PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-207370
(43)Date of publication of application : 28.07.2000

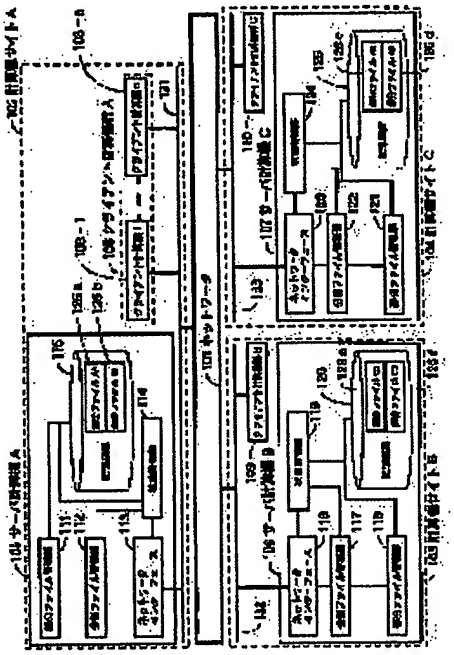
(51)Int.Cl. G06F 15/177
G06F 12/00
G06F 15/16

(21)Application number : 11-011513
(22)Date of filing : 20.01.1999

(71)Applicant : MATSUSHITA ELECTRIC IND CO LTD
(72)Inventor : SATO MASAKI
UESUGI AKIO
YASUKOCHI RYUJI
TANAKA NORIKO

(54) DISTRIBUTED FILE MANAGEMENT DEVICE AND DISTRIBUTED FILE MANAGEMENT SYSTEM

(57)Abstract:
PROBLEM TO BE SOLVED: To provide a distributed film management system which can make appropriate load distribution by means of plural server computers for generating, referring to and updating files.
SOLUTION: A distributed file management system is provided with server computers A, B, and C, client computer groups 108–110, and a network 101. The server computer A 105 is constituted of a storage device 115 which records partial files, a network interface 113, a partial file management section 111 which controls the write and read of the partial files, a status management section 114 which holds load information, and a distributed file management section 112. Since the arrangement of the partial files is determined, based on the load information of each server computer A, B, and C, the concentration of loads to a specific server computer can be avoided.



[0159] (Embodiment 5) Fig. 20 is a configuration diagram showing one example of the fifth embodiment of the distributed file management system of the present invention. In Fig. 20, the same codes are used for the similar configurations in Fig. 8. The distributed file management system shown in Fig. 20 comprises: a plurality of computer sites A2002, B2003, and C2004 comprising a client computer group composed of server computers such as a personal computer and a workstation and a plurality of client computers such as a personal computer and a workstation; and a network 101 such as a local area network and a wide area network mutually connecting the computer sites A2002, B2003, and C2004.

[0160] Here, the computer site A2002 comprises a plurality of server computers (only a "server computer A2005" is shown in Fig. 20) such as a personal computer and a workstation and a client computer group A108 composed of client computers 1 to n (108-1 to 108-n) such as a personal computer and a workstation. This computer site A2002 connects a plurality of server computers (only the "server computer A2005" is shown in Fig. 20) with the client computer group A108 in an internal network 131 such as Ethernet. The computer site A2002 works as, for example, an internet domain.

[0161] Similarly as with the computer site A2002, the computer site B2003 comprises a plurality of server computers (only a "server computer B2006" is shown in Fig. 20) and a client computer group B109 composed of a plurality of client computers, and the computer site C2004 comprises a plurality of server computers (only a "server computer C2007" is shown in Fig. 20) and a client computer group C110 composed of a plurality of client computers. Similarly as with the computer site A2002, these computer sites B2003 and C2004 further connect a plurality of server computers (only the "server computer B2006" and the "server computer C2007" are shown in Fig. 20) with the client computer groups B109 and C110 respectively in internal networks 132 and 133, and the computer sites B2003 and C2004 work as an internet domain.

[0162] The server computer A2005 is composed of a memory device 115 such as a hard disk recording a partial file of a distributed file, a network interface 113 for connecting to the internal network 131 such as Ethernet, a partial file management unit 111 controlling the writing and readout of the memory device 115 that records a partial file, a state management unit 814 monitoring the load to the memory device 115, the residual capacity of the memory device 115, and the load to the network interface 113 and maintaining information with regard to these loads and capacities, and a distributed file management unit 2012 connected to the partial file management unit 111, the state management unit 814, and the network interface 113.

[0163] This state management unit 814 notifies load information to other server

computers and comprises an external site management unit 811 maintaining external load information notified from another server computer.

[0164] The distributed file management unit 2012 determines, on the basis of each piece of information of each partial file obtained from an access information table 1201 (Fig. 12), a load information table 401 (Fig. 4), and an external load information table 901 (Fig. 9), a partial file to copy, and comprises a distributed file copy unit 2001 copying the partial file for another server computer.

[0165] The server computers B2006 and C2007 are configured similarly as with the server computer A2005. In other words, server computer B2006 is configured by a memory device 120, a network interface 118, a partial file management unit 116, a state management unit 819 comprising an external site management unit 812, and a distributed file management unit 2017 comprising a distributed file copy unit 2032. The server computer C2007 is configured by a memory device 125, a network interface 123, a partial file management unit 121, a state management unit 824 comprising an external site management unit 813, and a distributed file management unit 2022 comprising a distributed file copy unit 2033.

[0166] Here, the difference between the distributed file management systems shown in Figs. 20 and 8 is that the distributed file management units 2012, 2017, and 2022 shown in Fig. 20 determines, on the basis of each piece of information of each partial file obtained from the access information table 1201, the load information table 401, and the external load information table 901, a partial file to copy and comprises the distributed file copy units 2031, 2032, and 2033 copying the partial file for another server computer.

[0167] The operation of the distributed file management system configured as in the above will be described in detail, with the case being regarded as being an example in which, after distributed files A, B, and C are generated by the server computer A2005 as shown in Fig. 20, a partial file is copied.

[0168] Fig. 21 is a flowchart showing an operation algorithm of the distributed file copy unit 2031 of the server computer A2005. In Fig. 21, the distributed file copy unit 2031 firstly monitors, at prescribed time intervals, the load information table 401 (Fig. 4) that the state management unit 814 manages (step 2101).

[0169] The distributed file copy unit 2031 detects a "load" of any memory unit of the memory device 115 exceeding a prescribed value, such as 80 [%] for example (step 2101), and refers to a partial file management table 1301 (Fig. 13) to locate the partial file contained in a memory unit of the detected memory device 115. Then, the distributed file copy unit 2031 obtains the located access information of the located partial file from the access information table 1201. The distributed file copy unit 2031 compares the

"number of accesses" of the obtained access information and selects a partial file having the biggest "number of accesses" as a copy source partial file. Here, the situation will be considered in which a partial file A1 (126a) is, for example, selected. Now, on the basis of the external load information table 901 (Fig. 9), a server computer having a memory device having a sufficient "residual capacity" and a "load" being lower than a prescribed value of a memory unit of the memory device is selected. Then, it will be confirmed whether or not the partial file can be copied for the selected server computer and the server computer for which the partial file can be copied is selected as a copy destination server computer (step 2102).

[0170] Here, the situation will be considered in which a memory unit (memory device identifier: Disk ID2) of the memory device 125 of the server computer C2007 is, for example, selected.

[0171] On the basis of the information obtained according to the selection of this copy source partial file and the selection of the copy destination server computer, the partial file is copied (step 2103), and the monitoring processing of step 2101 is continued again.

[0172] Here, in the case of the example above, since the partial file A1 (126a) was selected as a copy source partial file and the server computer C2007 was selected as a copy destination server computer, the distributed file copy unit 2031 of the server computer A2005 being a copy source reads the partial file A1 (126a) out of the memory device 115 via the partial file management unit 111, and transmits this partial file A1 (126a) to the network 101 via the network interface 113. Additionally, at the same time, the information with regard to an "original location" (Fig. 13) of the partial file A1 (126a) is also transmitted.

[0173] Meanwhile, in the server computer C2007 being a copy destination, the distributed file copy unit 2033 receives the partial file A1 (126a) transmitted from the server computer A2005 being the copy source via the network interface 123. Then, the distributed file copy unit 2033 writes the partial file A1 (126a) to a prescribed memory unit of the memory device 125. In addition, the distributed file copy unit 2033 also receives the "original location" of the partial file A1 (126a) and registers it in the partial file management table 1403. After this, the server computer C2007 being the copy source notifies the completion of copying the partial file A1 (126a) to the computer A2005 being the copy source and the server computer shown in the "original location" (in this example, the "original location" is also a memory unit of the memory device 115 of the server computer A2005). The information of the partial file A1 (126a) of the partial file management table 1401 is rewritten in the server computer being the copy source and the server computer shown in the "original location" (both are the server

computer A2005).

[0174] Fig. 22 is a diagram showing a partial file management table of a server computer after copy processing. As a result of copying the partial file A1 (126a) above, the partial file management tables 1401, 1402, and 1403 shown in Fig. 14 are changed to be in the condition of the partial file management tables 2201, 2202, and 2203 shown in Fig. 22. In other words, Fig. 22 (A) shows the partial file management table 2201 of the server computer A2005, (B) shows the partial file management table 2202 of the server computer B2006, and (C) shows the partial file management table 2203 of the server computer C2007. In the partial file management table 2201, are shown a "location" and an "original location" of each of the partial files having "partial file identifiers" shown by A1, A2, A3, B1, C1, and C2. In the partial file management table 2002, are shown a "location" and an "original location" of each of the partial files having "partial file identifiers" shown by C1 and C2. In the partial file management table 2203, are shown a "location" and an "original location" of each of the partial files having "partial file identifiers" shown by A1, A2, and A3. The difference between the condition of the partial file management tables shown in Fig. 14 and the condition of the partial file management tables shown in Fig. 22 is the difference made by copying the partial file A1 (126a) from the server computer A2005 into the server computer C2007. In other words, the condition of the partial file management tables shown in Fig. 14 and the condition of the partial file management tables shown in Fig. 22 are different in that the "location" of the partial file A1 (126a) is "file://siteA/serverA/DisilD1/" in the partial file management table 1401 of Fig. (A), while the "locations" of the partial file A1 (126a) are "file://siteA/serverA/DiskID1/" and "file://siteC/serverC/DiskID2/ (the location of the copy)" in the partial file management table 2201 of the Fig. 22 (A). In the partial file management table 2103 of Fig. 22 (C), the item of the partial file A1 (126a) is further added.

[0175] According to the condition of Fig. 22, the distributed file copy unit 2032 of the server computer B2006 can copy the partial file C1 (126e) into the server computer C2007, and the distributed file copy unit 2033 of the server computer C2007 can also copy the partial file A1 (126a) into the server computer B2006.

[0176] Fig. 23 shows a diagram showing the partial file management tables in the condition of Fig. 22 with the partial file having been further copied. Fig. 23 (A) shows a partial file management table 2301 of the server computer A2005, (B) shows a partial file management table 2302 of the server computer B2006, and (C) shows a partial file management table 2303 of the server computer C2007. In the partial file management table 2301 of Fig. 23 (A), are shown a "location" and an "original location" of each of the

partial files having "partial file identifiers" being A1, A2, A3, B1, C1, and C2. In the partial file management table 2302 of Fig. 23 (B), are shown a "location" and an "original location" of each of the partial files having "partial file identifiers" being C1, C2, and A1. In the partial file management table 2303 of Fig. 23 (C), are shown a "location" and an "original location" of each of the partial files having "partial file identifiers" being A1, A2, A3, and C1. The difference between the condition of the partial file management tables in Fig. 22 and the condition of the partial file management tables shown in Fig. 23 is made by copying the partial file A1 (126a) from the server computer C2007 into the server computer B2006 and copying the partial file C1 (126e) from the server computer B2006 into the server computer C2007.

[0177] In other words, in correspondence with the copy of the partial file A1 (126a), the item of the partial file A1 (126a) is added in the partial file management table 2302. In addition, by the server computer B2006 referring to the "original location" of the partial file A1 (126a) to notify the server computer A2005 of a copy, the server computer A2005 of the "original location" adds "file://siteB/serverB/DiskID2/" to the "location" of the partial file A1 (126a) in the partial file management table 2301. In addition, in correspondence with the copy of the partial file C1 (126e), the item of the partial file C1 (126e) is added in the partial file management table 2303. Additionally, in the partial file management table 2302, the "locations" of the partial file C1 (126e) are "file://siteB/serverB/DiskID3/" and "file://siteC/serverC/DiskID3/ (the location of the copy)".

[0178] Described will be the operation in the condition shown in Fig. 23 in the case in which, when the client computer 1 (108-1) refers to the distributed file C, the reference contents are contained in the partial file C1 (126e).

[0179] (1) The client computer 1 (108-1) requests the reference of the distributed file C to the server computer A2005 that generated the distributed file C. In the server computer A2005, the distribution management table 2301 is referred to in order to inspect which partial file between the partial files C1 and C2 that configure the distributed file C is being referred to, and it is recognized that the partial file C1 (126e) is being referred to. Since the "location" of the partial file C1 (126e) is "file://siteB/serverB/DiskID3/" in the distributed file management table 2301, the server computer A2005 confirms whether or not the partial file C1 (126e) exists in the server computer B2006.

[0180] (2) The server computer B2006 inspects the distributed file management table 2302 to inspect the location of the partial file C1 (126e). Since the "locations" of the partial file C1 (126e) are "file://siteB/serverB/DiskID3/" and "file:

//siteC/serverC/DiskID3/" in the distributed file table 2302, the server computer B2006 selects a server computer having a low load between the server computers B2006 and C2007 from the load information table 401 and the external load information table 901. Here, when the server computer C2007 is selected, the server computer B2006 confirms whether or not the partial file C1 (126e) exists in the server computer C2007.

[0181] (3) The server computer C2007 inspects the distributed file management table 2303 to confirm the "location" of the partial file C1 (126e). Since the "location" of the partial file C1 (126e) is "file://siteC/serverC/DiskID3/" in the distributed file management table 2302, the partial file C1 (126e) is found out to exist in the server computer C2007.

[0182] (4) The server computer C2007 notifies to the server computer B2006 that the partial file C1 (126e) exists in "file://siteC/serverC/DiskID3".

[0183] (5) The server computer B2006 notifies to the server computer A2005 that the partial file C1 (126e) exists in "file://siteC/serverC/DiskID3".

[0184] (6) The server computer A2005 requests the reference of the partial file C1 (126e) to the server computer C2700. At the same time as this request, the client computer 1 (108-1) directly directs the server computer C2007 to perform a reference request of the partial file C1 (126e) for the client computer 1 (108-1) that requested the reference.

[0185] As is in the above, in the present embodiment, distributed file management units 112, 117, and 122 determine, on the basis of each piece of information of each partial file obtained from the access information table 1201, the load information table 401, and the external load information table 901, a partial file to copy, and the distributed file copy units 2031, 2032, and 2033 copies the partial file into another server computer so that loads can be prevented from concentrating on a memory device of a specific server computer.

[0186] In the fifth embodiment described above, in step 2101 of the operation algorithm of the distributed file copy unit 2031 (Fig. 21), instead of detecting a "load" of each memory unit of a memory device exceeding a prescribed value, a link having a "used communication bandwidth" of a network load information table 403 exceeding a prescribed value may be detected. In addition, in step 2102, a partial file strengthening a network load and a computer site strengthening a network load may be selected from the access information table 1201. This makes it possible to avoid concentrated network loads. When, for example, the "used communication bandwidth" of a link of the computer site A2002 (source site) and computer site B2003 (destination site) exceeds a prescribed value, the partial file existing in the server computer A2005

and causing a network load to be strengthen is copied into the server computer B2003. This reduces the "used communication bandwidth" between the computer sites A2002 and B2003.

[0187] Fig. 24 is a flowchart showing another operation algorithm of the distributed file copy unit 2031 of the server computer A2005. In Fig. 24, the distributed file copy unit 2031 firstly monitors a communication cost to a partial file at prescribed time intervals (step 2401).

[0188] Here, the communication cost can be, for example, communication time between a client computer referring to a partial file and a server computer maintaining the partial file. In, for example, the case of the communication cost of the client computer 1 (108-1) and the partial file A2 (126c), the communication cost can be the communication time between the client computer 1 (108-1) referring to the partial file A2 (126c) and the server computer C2007 maintaining the partial file A2 (126c).

[0189] Next, when the distributed file copy unit 2031 detects a communication cost to a partial file exceeding a prescribed value (step 2401), the distributed file copy unit 2031 selects as a copy source partial file the partial file that has this communication cost exceeding the prescribed value. When a plurality of client computers are accessing the partial file having a communication cost exceeding a prescribed value, each communication cost of each of the accesses is calculated and these communication costs are added to calculated a total communication cost. Meanwhile, when a server computer being a copy destination is selected, on the basis of the external load information table 901, the server computer having a sufficient "residual capacity" and a low "load" of a memory unit of a memory device is selected, a total communication cost is transmitted to the selected server computer, and how a communication cost changes as a result of a partial file being copied is sequentially inquired. Then, the server computer having the smallest communication cost is selected as a copy destination server computer. Alternatively, the distributed file copy units 2031, 2032, and 2033 may have, for example, a connection information table 1901 shown in Fig. 19 as connection information between sites (between a server computer and client computer), and by predicting a communication cost (communication time) according to the information of the connection information table 1901, the server computer having the smallest communication cost may be selected. According to the information of the connection information table 1901 in Fig. 19 and the information which site the server computer belongs to and which site the client computer belongs to, the "communication time" between the server computer and the client computer can be obtained. Then, the server computer having the smallest "communication time" can be selected as a copy

destination of the partial file. When a plurality of client computers are accessing the partial file having a communication cost exceeding a prescribed value, the server computer having the smallest total communication cost is selected as a copy destination server computer (step 2402).

[0190] When, for example, the communication cost (cost A2) of the client computer 1 (108-1) and the partial file A2 exceeds a prescribed value, the server computer A105 is inquired how the communication cost will change in the case of the partial file A2 being copied into the server computer A105. Alternatively, the cost is calculated by connection information 1801. If the result is less that the cost A2, the server computer A105 will be a candidate to be a copy destination. By repeating the above processing, the server computer having the smallest communication cost is searched for. On the basis of this piece of information obtained by the selection of a copy source partial file and the selection of a copy destination server computer, the partial file is copied (step 2403), and the monitoring processing of step 2401 is continued again.

[0191] As is in the above, by changing steps 2101 and 2102 in Fig. 21 to steps 2401 and 2402 in Fig. 24, the mean value of access time from a client computer to a partial file can be shortened. Although, in the above, "communication time" is taken for example as a communication cost, the communication cost may be "delay" or "fluctuation (fluctuation range)" of communication time.

[0192] In step 2102 in Fig. 21 and in step 2402 in Fig. 24, it is verified whether or not a partial file can be copied into a server computer being a copy destination, and in steps 2103 and 2403, the partial file is copied, however by copying the partial file without confirming whether or not the partial file can be copied into the server computer in steps 2102 and 2402, the confirmation processing in steps 2102 and 2402 can be omitted. In this case, if the copy of the partial file cannot be accepted in the server computer side being the copy destination, the server computer being the copy destination may further search for a copy destination for copying the partial file, or discard the partial file transmitted to be copied in order to notify to the server computer being a copy source that the partial file to copy was discarded.

[0193] In addition, in the fifth embodiment above, a partial file is merely copied form the server computer being a copy source into the server computer being a copy destination in step 2103 of Fig. 2 and in step 2403 of Fig. 24. In addition to the copy processing above, the partial file being possible to be moved to the server computer being the copy source can be selected from among partial files in the server computer being the copy destination, and the partial file can be moved to the server computer being the copy source. According to this, partial files are not concentrated on a single

server computer and loads concentrating on a memory device of a specific server computer can be avoided.

[0194] In step 2102 of Fig. 21, when a server computer being a copy destination is selected, it is also possible to set a server computer list in advance so that a server computer having a memory device having a sufficient residual capacity and a low load of a memory unit of the memory device can be selected from the server computers in this server computer list. This makes it possible to shorten the time to select a server.

(19)日本国特許庁（ＪＰ）　(12)**公開特許公報**（Ａ）　(11)特許出願公開番号
特開2000－207370
（P2000－207370A）
(43)公開日　平成12年７月28日(2000.7.28)

(51)Int.Cl.⁷		識別記号	FI			テーマコード（参考）
Ｇ０６Ｆ	15/177	６７４	Ｇ０６Ｆ	15/177	６７４Ａ	５Ｂ０４５
	12/00	５４５		12/00	５４５Ｂ	５Ｂ０８２
	15/16	６２０		15/16	６２０Ｂ	

審査請求　未請求　請求項の数31　ＯＬ　（全 42 頁）

(21)出願番号　特願平11－11513

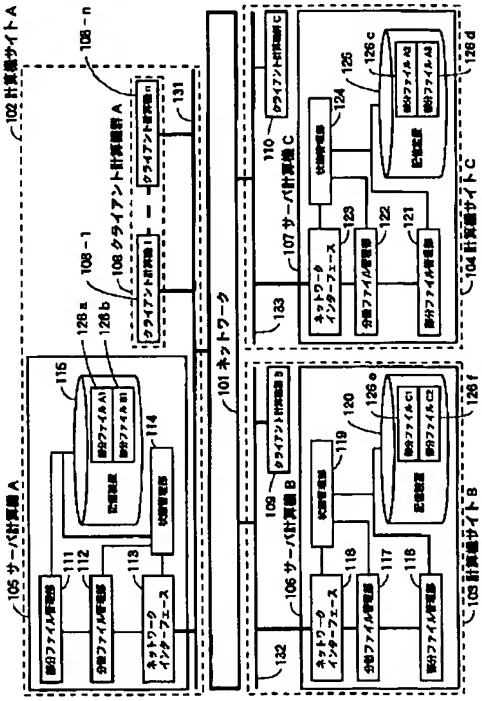
(22)出願日　平成11年１月20日(1999.1.20)

(71)出願人　000005821
松下電器産業株式会社
大阪府門真市大字門真1006番地
(72)発明者　佐藤　正樹
大阪府門真市大字門真1006番地　松下電器産業株式会社内
(72)発明者　上杉　明夫
大阪府門真市大字門真1006番地　松下電器産業株式会社内
(74)代理人　100107526
弁理士　鈴木　直郁　（外１名）

最終頁に続く

(54)【発明の名称】　分散ファイル管理装置及び分散ファイル管理システム

(57)【要約】
【課題】　ファイルの作成、参照、更新において、複数のサーバ計算機で適切な負荷分散を行うことができる分散ファイル管理システムを提供する。
【解決手段】　本発明の分散ファイル管理システムは、サーバ計算機Ａ、Ｂ、Ｃと、クライアント計算機群１０８～１１０と、ネットワーク１０１とを備えている。サーバ計算機Ａ１０５は、部分ファイルを記録する記憶装置１１５と、ネットワークインタフェース１１３と、部分ファイルの書き込みや読み出しを制御する部分ファイル管理部１１１と、負荷を監視し、負荷情報を保持する状態管理部１１４と、分散ファイル管理部１１２とによって構成されている。各サーバ計算機Ａ、Ｂ、Ｃの負荷情報に基づいて部分ファイルの配置を決定するため、特定のサーバ計算機への負荷の集中を回避することができる



1

【特許請求の範囲】

【請求項１】データを記憶する記憶手段を有する複数のサーバ計算機と１または複数のクライアント計算機とが接続されているネットワークに接続された分散ファイル管理装置であって、
前記複数のサーバ計算機の負荷情報を保持して管理する状態管理手段と、
前記クライアント計算機からの分散ファイルの処理要求に対応して、前記分散ファイルの部分ファイルを特定し、前記状態管理手段で管理されている前記負荷情報に基づいて、前記部分ファイルを処理するサーバ計算機を決定する分散ファイル管理手段と、
を備える、ことを特徴とする分散ファイル管理装置。

【請求項２】前記状態管理手段は、他の分散ファイル管理装置へ前記負荷情報を通知し、また、他の分散ファイル管理装置から通知されたサーバ計算機の負荷情報を外部負荷情報として保持する外部状態管理手段を備える、ことを特徴とする請求項１記載の分散ファイル管理装置。

【請求項３】前記外部状態管理手段は、マルチキャストによって前記他の分散ファイル管理装置へ前記負荷情報を通知する、ことを特徴とする請求項２記載の分散ファイル管理装置。

【請求項４】前記外部状態管理手段は、前記他の分散ファイル管理装置のうち、隣接する他の分散ファイル管理装置に前記負荷情報を通知する、ことを特徴とする請求項２または３記載の分散ファイル管理装置。

【請求項５】複数のサーバ計算機と、１または複数のクライアント計算機と、前記複数のサーバ計算機及び前記１または複数のクライアント計算機を接続するネットワークとを備えた分散ファイル管理システムにおいて、
前記複数のサーバ計算機の各々は、
分散ファイルの一部または全部を構成する部分ファイルを記憶する記憶手段と、
負荷情報を保持して管理する状態管理手段と、
前記クライアント計算機からの分散ファイルの処理要求に対応して、前記分散ファイルの部分ファイルを特定し、前記状態管理手段で管理されている前記負荷情報に基づいて、前記部分ファイルを処理するサーバ計算機を決定する分散ファイル管理手段と、
を備える、
ことを特徴とする分散ファイル管理システム。

【請求項６】前記状態管理手段は、他のサーバ計算機へ前記負荷情報を通知し、また、他のサーバ計算機から通知された当該他のサーバ計算機の負荷情報を外部負荷情報として保持する外部状態管理手段を備える、ことを特徴とする請求項５記載の分散ファイル管理システム。

【請求項７】前記外部状態管理手段は、マルチキャストによって前記他のサーバ計算機へ前記負荷情報を通知する、ことを特徴とする請求項６記載の分散ファイル管理システム。

2

【請求項８】前記複数のサーバ計算機は、１または複数のサーバ計算機群にグループ分けされており、
前記外部状態管理手段は、前記１または複数のサーバ計算機群のうち、所定のサーバ計算機群に属する他のサーバ計算機へ前記負荷情報を通知する、
ことを特徴とする請求項６または７記載の分散ファイル管理システム。

【請求項９】前記複数のサーバ計算機は、１または複数のサーバ計算機群にグループ分けされており、
前記外部状態管理手段は、前記１または複数のサーバ計算機群のうち、隣接するサーバ計算機群に属する他のサーバ計算機に前記負荷情報を通知する、
ことを特徴とする請求項６または７記載の分散ファイル管理システム。

【請求項１０】前記分散ファイル管理手段は、前記部分ファイル毎のアクセス情報、前記負荷情報、及び前記外部負荷情報に基づいて移動する部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該サーバ計算機へ移動する分散ファイル移動手段を備える、ことを特徴とする請求項２乃至４記載の分散ファイル管理装置または６乃至９記載の分散ファイル管理システム。

【請求項１１】前記分散ファイル移動手段は、前記負荷情報に含まれる前記記憶手段の負荷が所定の値よりも大であることを検知し、前記外部負荷情報と前記アクセス情報に基づいて移動する部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へ移動する、ことを特徴とする請求項１０記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項１２】前記分散ファイル移動手段は、前記負荷情報に含まれる前記記憶手段の残容量が所定の値より小であることを検知し、前記外部負荷情報と前記アクセス情報に基づいて移動する部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へ移動する、ことを特徴とする請求項１０記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項１３】前記分散ファイル移動手段は、前記負荷情報に含まれる前記ネットワークの負荷が所定の値よりも大であることを検知し、前記外部負荷情報と前記アクセス情報に基づいて移動する部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へ移動する、ことを特徴とする請求項１０記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項１４】前記分散ファイル移動手段は、前記負荷情報、前記外部負荷情報、前記アクセス情報、及び前記クライアント計算機と前記複数のサーバ計算機との間の接続情報に基づいて、前記部分ファイルを保持している記憶手段を有するサーバ計算機と処理要求を行ったクライアント計算機との間の通信コストを求め、該通信コス

トよりも小の通信コストとなる他のサーバ計算機を決定し、該他のサーバ計算機へ前記部分ファイルを移動する、ことを特徴とする請求項１０記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項１５】前記分散ファイル移動手段は、前記部分ファイルの移動先となる前記他のサーバ計算機に対して、予め前記部分ファイルの移動が可能か否かを確認する、ことを特徴とする請求項１０乃至１４記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項１６】前記分散ファイル移動手段は、前記部分ファイルを前記他のサーバ計算機へ移動した際に、前記他のサーバ計算機から前記サーバ計算機に他の部分ファイルを移動する、ことを特徴とする請求項１０乃至１５記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項１７】前記分散ファイル移動手段は、前記部分ファイルを移動することができる他のサーバ計算機の候補をリストにし、該リストに基づいて、前記部分ファイルを移動する他のサーバ計算機を決定する、ことを特徴とする請求項１０乃至１６記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項１８】前記分散ファイル移動手段は、前記部分ファイルの移動と共に、前記部分ファイルを作成したサーバ計算機に関する情報を前記他のサーバ計算機に送る、ことを特徴とする請求項１０乃至１７記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項１９】前記分散ファイル管理手段は、前記部分ファイル毎のアクセス情報、前記負荷情報、及び前記外部負荷情報に基づいて、コピーする部分ファイルとコピー先の他のサーバ計算機を決定し、前記部分ファイルを前記他のサーバ計算機にコピーする分散ファイルコピー手段を備える、ことを特徴とする請求項2乃至4、6乃至１８記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項２０】前記分散ファイルコピー手段は、前記負荷情報に含まれる前記記憶手段の負荷が所定の値よりも大であることを検知し、前記外部負荷情報と前記アクセス情報に基づいてコピーする部分ファイルとコピー先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へコピーする、ことを特徴とする請求項１９記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項２１】前記分散ファイルコピー手段は、前記負荷情報に含まれる前記ネットワークの負荷が所定の値よりも大であることを検知し、前記外部負荷情報と前記アクセス情報に基づいてコピーする部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へコピーする、ことを特徴とする請求項１９記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項２２】前記分散ファイルコピー手段は、前記負荷情報、前記外部負荷情報、前記アクセス情報、及び前記クライアント計算機と前記複数のサーバ計算機との間の接続情報に基づいて、前記部分ファイルを保持している記憶手段を有するサーバ計算機と処理要求を行ったクライアント計算機との間の通信コストを求め、該通信コストよりも小の通信コストとなる他のサーバ計算機を決定し、該他のサーバ計算機へ前記部分ファイルをコピーする、ことを特徴とする請求項１９記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項２３】前記分散ファイルコピー手段は、前記部分ファイルのコピー先となる前記他のサーバ計算機に対して、予め前記部分ファイルのコピーが可能か否かを確認する、ことを特徴とする請求項１９乃至２２記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項２４】前記分散ファイルコピー手段は、前記部分ファイルを前記他のサーバ計算機へコピーした際に、前記他のサーバ計算機から前記サーバ計算機に他の部分ファイルをコピーする、ことを特徴とする請求項１９乃至２３記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項２５】前記分散ファイルコピー手段は、前記部分ファイルをコピーすることができる他のサーバ計算機の候補をリストにし、該リストに基づいて、前記部分ファイルをコピーする他のサーバ計算機を決定する、ことを特徴とする請求項１９乃至２４記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項２６】前記分散ファイルコピー手段は、前記部分ファイルをコピーする複数のコピー先の他のサーバ計算機を選択し、選択された前記複数の他のサーバ計算機へマルチキャストによって同時に前記部分ファイルをコピーする、ことを特徴とする請求項１９乃至２５記載の分散ファイル管理装置及び分散ファイル管理システム。

【請求項２７】前記状態管理手段で管理されている前記負荷情報は、前記記憶手段の容量及び負荷、並びに前記ネットワークと前記複数のサーバ計算機との間の通信負荷を含む、ことを特徴とする請求項1乃至２６記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項２８】さらに、前記サーバ計算機は、前記部分ファイルを前記記憶手段に書き込み、また、前記部分ファイルを前記記憶手段から読み出す部分ファイル管理手段を備える、ことを特徴とする請求項1乃至２７記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項２９】前記分散ファイル管理手段は、
前記クライアント計算機からの前記処理要求が分散ファイルの作成要求の場合には、該分散ファイルを複数の部分ファイルに分割し、分割した部分ファイルを保持するサーバ計算機を前記状態管理手段で管理されている前記負荷情報に基づいて決定し、

前記クライアント計算機からの前記処理要求が分散ファイルの参照要求または更新要求の場合には、前記参照要求または前記更新要求の処理の対象となる部分ファイルの存在を決め、前記処理要求を処理するサーバ計算機を前記状態管理手段で管理されている前記負荷情報に基づいて決定する、

ことを特徴とする請求項１乃至２８記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項３０】前記分散ファイル管理手段は、前記クライアント計算機からの情報に基づいて、分散ファイルの一部または全部を構成する前記部分ファイルのサイズを決定する部分ファイルサイズ決定手段を備える、ことを特徴とする請求項１乃至２９記載の分散ファイル管理装置または分散ファイル管理システム。

【請求項３１】前記分散ファイル管理手段は、分散ファイルに記録されているデータの種類に基づいて、前記分散ファイルの一部または全部を構成する前記部分ファイルのサイズを決定する部分ファイルサイズ決定手段を備える、ことを特徴とする請求項１乃至２９記載の分散ファイル管理装置または分散ファイル管理システム。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】本発明は、コンピュータネットワークシステムにおいて複数の端末にファイルを分散して管理する分散ファイル管理装置及び分散ファイル管理システムに関する。特には、複数のサーバ計算機及びクライアント計算機をネットワークで接続したサーバ・クライアント型のコンピュータネットワークシステムにおいて、複数のサーバ計算機でファイルを分散して管理する分散ファイル管理装置及び分散ファイル管理システムに関する。

【０００２】

【従来の技術】従来から、複数のサーバ計算機及びクライアント計算機をネットワークで接続したサーバ・クライアント型のコンピュータネットワークシステム（以下、単に「ネットワークシステム」ともいう）に適用される分散ファイル管理技術として、例えば、特開平８－７７０５４号広報に開示されている分散ファイルシステムなどがある。

【０００３】図３０は、特開平８－７７０５４号広報に開示されている従来の分散ファイルシステムを示す。図３０において、この従来の分散ファイルシステムは、複数のサーバ計算機３００２、３００３、３００４、３００５と、複数のクライアント計算機３００６、３００７、３００８と、これら複数のサーバ計算機３００２、３００３、３００４、３００５及び複数のクライアント計算機３００６、３００７、３００８を接続するネットワーク３００１と、を備えている。

【０００４】ここで、サーバ計算機３００２には、分散ファイルＡの部分ファイルＡ－１（３００２－１）と分散ファイルＢの部分ファイルＢ－１（３００２－３）が保持されている。また、サーバ計算機３００３には、分散ファイルＡの部分ファイルＡ－２（３００３－１）と分散ファイルＢの部分ファイルＢ－２（３００３－２）が保持されている。また、サーバ計算機３００４には、分散ファイルＡの部分ファイルＡ－３（３００４－１）と分散ファイルＢの部分ファイルＢ－３（３００４－２）が保持されている。

【０００５】また、サーバ計算機３００５は、各サーバ計算機３００２、３００３、３００４に保持されているそれぞれの部分ファイルＡ－１～Ａ－３、Ｂ－１～Ｂ－３を管理する分散ファイル管理部３００５－２を備え、クライアント計算機３００６、３００７、３００８からの部分ファイルの参照要求または更新要求（以下、単に「参照／更新要求」ともいう）に対してその振り分けを行なうための参照／更新要求振り分け情報３００５－１を保持している。

【０００６】一方、クライアント計算機３００６は、分散ファイル作成要求に応じて分散ファイルの部分ファイルを作成する分散ファイル作成部３００６－１と、分散ファイルに対する更新要求に応じて、サーバ計算機３００５の参照／更新要求振り分け情報３００５－１に基づいて該分散ファイルの部分ファイルの所在を決定する更新要求振り分け部３００６－２と、分散ファイルに対する参照要求に応じて、サーバ計算機３００５の参照／更新要求振り分け情報３００５－１に基づいて該分散ファイルの部分ファイルの所在を決定する参照要求振り分け部３００６－３と、を備えている。なお、他のクライアント計算機３００７、３００８も同様の構成になっている。

【０００７】上述した従来の分散ファイルシステムによれば、例えば、クライアント計算機３００６の利用者が分散ファイルの作成を要求した場合、クライアント計算機３００６の分散ファイル作成部３００６－１は、予め決められている振り分け条件に基づいて当該分散ファイルの部分ファイルを作成するサーバ計算機を決定し、該サーバ計算機に分散ファイルの部分ファイルを作成する。そして、この部分ファイルの作成と同時に、どのサーバ計算機に部分ファイルを作成したかを表す参照／更新要求振り分け情報３００５－１を生成する。この参照／更新要求振り分け情報３００５－１は、ネットワーク３００１を介してサーバ計算機３００５に送信され、サーバ計算機３００５で保持される。

【０００８】また、例えば、クライアント計算機３００６の利用者が分散ファイルを参照または更新する参照／更新要求を行なった場合、まず、クライアント計算機３００６は、サーバ計算機３００５の分散ファイル管理部３００５－２に対して、該当する分散ファイルのオープン要求を行なう。サーバ計算機３００５の分散ファイル管理部３００５－２は、クライアント計算機３００６か

らの分散ファイルのオープン要求に対応して、クライアント計算機３００６へネットワーク３００１を介して当該分散ファイルに関する参照／更新要求振り分け情報３００５－１を送信する。クライアント計算機３００６の更新要求振り分け部３００６－２または参照要求振り分け部３００６－３は、サーバ計算機３００５から受け取った参照／更新要求振り分け情報３００５－１に基づいて、分散ファイルの部分ファイルを保持しているサーバ計算機に対し、参照／更新要求を送信する。

【０００９】このように、従来の分散ファイルシステムにおいては、分散ファイルを複数の部分ファイルに分割し、分散ファイルに対する処理（作成、参照、更新）を部分ファイル単位の処理に分散することにより、１つの分散ファイルに対して複数の処理要求が集中した場合にも、１つのサーバ計算機に負荷を集中させずに、負荷の分散を行なうことができる。

【００１０】

【発明が解決しようとする課題】しかしながら、図３０に示したような従来の分散ファイルシステムにおいては、分散ファイルを構成する部分ファイルを作成する際に、予め決められている固定的な振り分け規則に基づいて、該部分ファイルを作成するサーバ計算機を決定しているため、サーバ計算機へのファイルの分散において現実的なサーバ計算機の負荷情報を考慮したものにはなっていない。このため、実際には、アクセスや処理が集中して負荷の高くなっているサーバ計算機に対して、さらに部分ファイルの作成要求が発生することがある。そのため、特定のサーバ計算機のみの負荷が大きくなる場合があり、複数のサーバ計算機で適切な負荷分散が行なわれないという問題があった。

【００１１】また、従来の分散ファイルシステムにおいては、上記の振り分け規則に基づいて、一度固定的に部分ファイルの振り分けをサーバ計算機に行なってしまうと、その振り分け以後、すなわち、振り分けられたサーバ計算機での部分ファイルの作成以後には、作成された部分ファイルの移動やコピーを行なわないため、特定の部分ファイルへのアクセスが集中した場合、アクセスによる負荷の分散を行なうことができないという問題があった。

【００１２】したがって、本発明の目的は、ファイルの作成、参照、更新において、複数のサーバ計算機で適切な負荷分散を行うことができる分散ファイル管理装置及び分散ファイル管理システムを提供することである。

【００１３】

【課題を解決するための手段】上記課題を解決するために、本発明に係る第１の態様の分散ファイル管理装置は、　データを記憶する記憶手段を有する複数のサーバ計算機と１または複数のクライアント計算機とが接続されているネットワークに接続された分散ファイル管理装置であって、　複数のサーバ計算機の負荷情報を保持して管理する状態管理手段と、　クライアント計算機からの分散ファイルの処理要求に対応して、分散ファイルの部分ファイルを特定し、状態管理手段で管理されている負荷情報に基づいて、部分ファイルを処理するサーバ計算機を決定する分散ファイル管理手段と、を備えることを特徴とする。

【００１４】上述の本発明に係る分散ファイル管理装置において、　状態管理手段は、他の分散ファイル管理装置へ負荷情報を通知し、また、他の分散ファイル管理装置から通知されたサーバ計算機の負荷情報を外部負荷情報として保持する外部状態管理手段を備えることもできる。

【００１５】ここで、外部状態管理手段は、マルチキャストによって他の分散ファイル管理装置へ負荷情報を通知するようにしてもよい。また、外部状態管理手段は、他の分散ファイル管理装置のうち、隣接する他の分散ファイル管理装置に負荷情報を通知することもできる。

【００１６】また、上記課題を解決するために、本発明に係る第１の態様の分散ファイル管理システムは、　複数のサーバ計算機と、１または複数のクライアント計算機と、複数のサーバ計算機及び１または複数のクライアント計算機を接続するネットワークとを備えた分散ファイル管理システムにおいて、　複数のサーバ計算機の各々は、　分散ファイルの一部または全部を構成する部分ファイルを記憶する記憶手段と、　負荷情報を保持して管理する状態管理手段と、　クライアント計算機からの分散ファイルの処理要求に対応して、分散ファイルの部分ファイルを特定し、状態管理手段で管理されている負荷情報に基づいて、部分ファイルを処理するサーバ計算機を決定する分散ファイル管理手段と、を備えることを特徴とする。

【００１７】上述の本発明に係る分散ファイル管理システムにおいて、　状態管理手段は、他のサーバ計算機へ負荷情報を通知し、また、他のサーバ計算機から通知された当該他のサーバ計算機の負荷情報を外部負荷情報として保持する外部状態管理手段を備えるようにしてもよい。　ここで、外部状態管理手段は、マルチキャストによって他のサーバ計算機へ負荷情報を通知することもできる。

【００１８】また、上述の本発明に係る分散ファイル管理装置及び分散ファイル管理システムにおいて、　複数のサーバ計算機は、１または複数のサーバ計算機群にグループ分けされており、　外部状態管理手段は、１または複数のサーバ計算機群のうち、所定のサーバ計算機群に属する他のサーバ計算機へ負荷情報を通知するようにしてもよく、または、外部状態管理手段は、１または複数のサーバ計算機群のうち、隣接するサーバ計算機群に属する他のサーバ計算機に前記負荷情報を通知するようにしてもよい。

【００１９】さらに、分散ファイル管理手段は、部分フ

ファイル毎のアクセス情報、負荷情報、及び外部負荷情報に基づいて移動する部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該サーバ計算機へ移動する分散ファイル移動手段を備えることもできる。

【００２０】ここで、分散ファイル移動手段は、負荷情報に含まれる記憶手段の負荷が所定の値よりも大であることを検知し、外部負荷情報とアクセス情報に基づいて移動する部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へ移動するようにしてもよく、　または、分散ファイル移動手段は、負荷情報に含まれる記憶手段の残容量が所定の値より小であることを検知し、外部負荷情報とアクセス情報に基づいて移動する部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へ移動するようにしてもよい。　また、分散ファイル移動手段は、負荷情報に含まれるネットワークの負荷が所定の値よりも大であることを検知し、外部負荷情報とアクセス情報に基づいて移動する部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へ移動することもでき、　または、分散ファイル移動手段は、負荷情報、外部負荷情報、アクセス情報、及びクライアント計算機と複数のサーバ計算機との間の接続情報に基づいて、部分ファイルを保持している記憶手段を有するサーバ計算機と処理要求を行ったクライアント計算機との間の通信コストを求め、該通信コストよりも小の通信コストとなる他のサーバ計算機を決定し、該他のサーバ計算機へ部分ファイルを移動するようにしてもよい。

【００２１】さらに、分散ファイル移動手段は、部分ファイルの移動先となる他のサーバ計算機に対して、予め前記部分ファイルの移動が可能か否かを確認することもでき、　また、分散ファイル移動手段は、部分ファイルを他のサーバ計算機へ移動した際に、他のサーバ計算機からサーバ計算機に他の部分ファイルを移動することもできる。　さらに、分散ファイル移動手段は、部分ファイルを移動することができる他のサーバ計算機の候補をリストにし、該リストに基づいて、部分ファイルを移動する他のサーバ計算機を決定するようにしてもよい。

また、分散ファイル移動手段は、部分ファイルの移動と共に、部分ファイルを作成したサーバ計算機に関する情報を他のサーバ計算機に送るようにすることもできる。

【００２２】また、上述の分散ファイル管理手段は、部分ファイル毎のアクセス情報、負荷情報、及び外部負荷情報に基づいて、コピーする部分ファイルとコピー先の他のサーバ計算機を決定し、部分ファイルを他のサーバ計算機にコピーする分散ファイルコピー手段を備えることもできる。

【００２３】このとき、分散ファイルコピー手段は、負荷情報に含まれる記憶手段の負荷が所定の値よりも大であることを検知し、外部負荷情報とアクセス情報に基づいてコピーする部分ファイルとコピー先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へコピーするようにしてもよく、　分散ファイルコピー手段は、負荷情報に含まれるネットワークの負荷が所定の値よりも大であることを検知し、外部負荷情報とアクセス情報に基づいてコピーする部分ファイルと移動先の他のサーバ計算機を決定し、該部分ファイルを該他のサーバ計算機へコピーするようにしてもよい。　または、分散ファイルコピー手段は、負荷情報、外部負荷情報、アクセス情報、及びクライアント計算機と複数のサーバ計算機との間の接続情報に基づいて、部分ファイルを保持している記憶手段を有するサーバ計算機と処理要求を行ったクライアント計算機との間の通信コストを求め、該通信コストよりも小の通信コストとなる他のサーバ計算機を決定し、該他のサーバ計算機へ部分ファイルをコピーするようにすることもできる。

【００２４】ここで、分散ファイルコピー手段は、部分ファイルのコピー先となる他のサーバ計算機に対して、予め部分ファイルのコピーが可能か否かを確認するようにするとよい。　また、分散ファイルコピー手段は、部分ファイルを他のサーバ計算機へコピーした際に、他のサーバ計算機からサーバ計算機に他の部分ファイルをコピーしてもよい。　また、分散ファイルコピー手段は、部分ファイルをコピーすることができる他のサーバ計算機の候補をリストにし、該リストに基づいて、部分ファイルをコピーする他のサーバ計算機を決定するようにすることもできる。　さらに、分散ファイルコピー手段は、部分ファイルをコピーする複数のコピー先の他のサーバ計算機を選択し、選択された複数の他のサーバ計算機へマルチキャストによって同時に部分ファイルをコピーすることもできる。

【００２５】また、状態管理手段で管理されている負荷情報は、記憶手段の容量及び負荷、並びにネットワークと複数のサーバ計算機との間の通信負荷を含むことができる。

【００２６】さらに、上述の本発明に係る分散ファイル管理装置及び分散ファイル管理システムにおいて、　サーバ計算機は、部分ファイルを記憶手段に書き込み、また、部分ファイルを記憶手段から読み出す部分ファイル管理手段を備えることもできる。

【００２７】また、分散ファイル管理手段は、　クライアント計算機からの処理要求が分散ファイルの作成要求の場合には、該分散ファイルを複数の部分ファイルに分割し、分割した部分ファイルを保持するサーバ計算機を状態管理手段で管理されている負荷情報に基づいて決定し、　クライアント計算機からの処理要求が分散ファイルの参照要求または更新要求の場合には、参照要求または更新要求の処理の対象となる部分ファイルの存在を決め、処理要求を処理するサーバ計算機を状態管理手段で管理されている負荷情報に基づいて決定する、ようにし

てもよい。

【００２８】また、分散ファイル管理手段は、クライアント計算機からの情報に基づいて、分散ファイルの一部または全部を構成する部分ファイルのサイズを決定する部分ファイルサイズ決定手段を備えるようにすることができ、　または、分散ファイル管理手段は、分散ファイルに記録されているデータの種類に基づいて、分散ファイルの一部または全部を構成する部分ファイルのサイズを決定する部分ファイルサイズ決定手段を備えるようにしてもよい。

【００２９】上述の本発明に係る分散ファイル管理装置及び分散ファイル管理システムにおいては、分散ファイル管理手段が、サーバ計算機の負荷情報に基づいて、部分ファイルを配置するサーバ計算機を決定するため、特定のサーバ計算機への負荷の集中を回避することができる。

【００３０】また、分散ファイル管理手段が、他のサーバ計算機の負荷情報に基づいて、部分ファイルを配置するサーバ計算機を決定するため、特定のサーバ計算機で負荷が集中することを回避できる。

【００３１】また、部分ファイルを他のサーバ計算機に移動することによって、特定のサーバ計算機の記憶手段への負荷の集中や、記憶手段の容量の不均衡を回避することができる。また、部分ファイルを他のサーバ計算機にコピーすることによって、特定のサーバ計算機の記憶装置への負荷の集中を回避することができる。

【００３２】また、分散ファイルを構成する部分ファイルのサイズを適宜変更することができるため、論理的、内容的に関連のあるデータ、例えば、画像１フレーム分のデータなどを複数の部分ファイルに分割してしまうことを避けられる。

【００３３】

【発明の実施の形態】以下、本発明の分散ファイル管理装置及び分散ファイル管理システムの実施の形態について、図１から図２９を用いて説明する。

【００３４】（実施の形態１）図１は、本発明における分散ファイル管理システムの第１の実施の形態の一例を示す構成図である。図１において、この分散ファイル管理システムは、パーソナルコンピュータやワークステーションなどのサーバ計算機及びパーソナルコンピュータやワークステーションなどの複数のクライアント計算機から成るクライアント計算機群を備えた複数の計算機サイトＡ１０２、計算機サイトＢ１０３、及び計算機サイトＣ１０４と、計算機サイトＡ１０２、計算機サイトＢ１０３、及び計算機サイトＣ１０４を相互に接続するローカルエリアネットワークやワイドエリアネットワークなどのネットワーク１０１とを備えている。

【００３５】ここで、計算機サイトＡ１０２は、パーソナルコンピュータやワークステーションなどの複数のサーバ計算機（図１においては、「サーバ計算機Ａ１０５」のみ示す）と、パーソナルコンピュータやワークステーションなどのクライアント計算機１～ｎ（１０８－１～１０８－ｎ）から成るクライアント計算機群Ａ１０８とを備えている。この計算機サイトＡ１０２は、複数のサーバ計算機（図１においては、「サーバ計算機Ａ１０５」のみ示す）とクライアント計算機群Ａ１０８とをイーサネットなどの内部ネットワーク１３１で接続しており、例えば、インターネットドメインになっている。

【００３６】また、計算機サイトＡ１０２と同様に、計算機サイトＢ１０３は、複数のサーバ計算機（図１においては、「サーバ計算機Ｂ１０６」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｂ１０９とを備え、計算機サイトＣ１０４は、複数のサーバ計算機（図１においては、「サーバ計算機Ｃ１０７」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｃ１１０とを備えている。さらに、これらの計算機サイトＢ１０３及び計算機サイトＣ１０４は、計算機サイトＡ１０２と同様に、複数のサーバ計算機（図１においては、「サーバ計算機Ｂ１０６」及び「サーバ計算機Ｃ１０７」のみ示す）と、クライアント計算機群Ｂ１０９及びクライアント計算機群Ｃ１１０とを、それぞれ内部ネットワーク１３２及び内部ネットワーク１３３で接続しており、例えば、インターネットドメインになっている。

【００３７】サーバ計算機Ａ１０５は、分散ファイルの部分ファイルを記録するハードディスクなどの記憶装置１１５と、イーサネットなどの内部ネットワーク１３１へ接続するためのネットワークインタフェース１１３と、部分ファイルを記録している記憶装置１１５への書き込みや読み出しを制御する部分ファイル管理部１１１と、記憶装置１１５に対する負荷や記憶装置１１５の残り容量、及びネットワークインタフェース１１３に対する負荷を監視し、これらの負荷や容量に関する情報を保持する状態管理部１１４と、部分ファイル管理部１１１、状態管理部１１４、及びネットワークインタフェース１１３に接続された分散ファイル管理部１１２とによって構成されている。

【００３８】この分散ファイル管理部１１２は、部分ファイルの書き込みや読み出しを部分ファイル管理部１１１に指示する。また、分散ファイル管理部１１２は、分散ファイルを作成する場合には、状態管理部１１４から得られる情報に基づいて、分散ファイルを複数の部分ファイルに分割し、各部分ファイルを配置（記録）するサーバ計算機を決定する。また、以前に作成された分散ファイルを参照または更新する場合には、該当する分散ファイルの部分ファイルが存在する（記録されている）サーバ計算機を決定する。

【００３９】サーバ計算機Ｂ１０６及びサーバ計算機Ｃ１０７は、サーバ計算機Ａ１０５と同様の構成になっている。すなわち、サーバ計算機Ｂ１０６は、記憶装置１

２０と、ネットワークインタフェース１１８と、部分ファイル管理部１１６と、状態管理部１１９と、分散ファイル管理部１１７とによって構成されている。また、サーバ計算機Ｃ１０７は、記憶装置１２５と、ネットワークインタフェース１２３と、部分ファイル管理部１２１と、状態管理部１２４と、分散ファイル管理部１２２とによって構成されている。

【００４０】図２は、分散ファイルの構成の一例を示す図である。図２において、分散ファイル２０１は、複数の部分ファイル２０２－１～２０２－ｎによって構成されている。

【００４１】図１においては、各分散ファイルＡ、Ｂ、Ｃが作成された後の状態を示している。すなわち、サーバ計算機Ａ１０５の記憶装置１１５には、分散ファイルＡの部分ファイルＡ１（１２６ａ）と分散ファイルＢの部分ファイルＢ１（１２６ｂ）とが記録されている。また、サーバ計算機Ｂ１０６の記憶装置１２０には、分散ファイルＣの部分ファイルＣ１（１２６ｅ）と分散ファイルＣの部分ファイルＣ２（１２６ｆ）とが記録されている。また、サーバ計算機Ｃ１０７の記憶装置１２５には、分散ファイルＡの部分ファイルＡ２（１２６ｃ）と分散ファイルＡの部分ファイルＡ３（１２６ｄ）とが記録されている。

【００４２】次に、以上のように構成された分散ファイル管理システムの動作について説明する。以下においては、クライアント計算機群Ａ１０８のクライアント計算機１（１０８－１）からサーバ計算機Ａ１０５に対して分散ファイルＡの作成要求が発行され、図１に示したような部分ファイルＡ１～Ａ３が作成される場合の分散処理を例にして説明する。ここで、図１に示した記憶装置１１５、１２０、１２５は、それぞれ複数の記憶部または記憶領域（以下、単に「記憶部」ともいう）を有するものとする。これらの複数の記憶部は、物理的に１つの記録媒体であってもよく、また、複数の記録媒体であってもよい。

【００４３】図１において、まず、クライアント計算機１（１０８－１）からサーバ計算機Ａ１０５に分散ファイルＡの作成要求が発行される。この分散ファイルＡの作成要求は、内部ネットワーク１３１及びサーバ計算機Ａ１０５のネットワークインタフェース１１３を介して、サーバ計算機Ａ１０５の分散ファイル管理部１１２によって受け取られる。

【００４４】図３は、分散ファイルの作成要求を受け取った場合の分散ファイル管理部の動作アルゴリズムを示すフローチャートである。以下、図１及び図３を用いて、分散ファイル管理部１１２の詳細な動作を説明する。

【００４５】分散ファイルＡの作成要求を受け取った分散ファイル管理部１１２は、まず、状態管理部１１４の管理している負荷情報を獲得する（ステップ３０１）。

【００４６】図４は、状態管理部１１４で管理する負荷情報テーブル４０１を示す図である。図４において、負荷情報テーブル４０１は、記憶装置負荷情報テーブル４０２及びネットワーク負荷情報テーブル４０３から成っている。

【００４７】この記憶装置負荷情報テーブル４０２は、サーバ計算機Ａ１０５に接続されている記憶装置１１５の複数の記憶部を識別するための「記憶装置識別子」と、各記憶部の負荷情報［％］を示す「負荷」と、各記憶部の残容量［Ｍbytes］を示す「残容量」の項目で構成されている。ここで、記憶装置１１５の各記憶部の「負荷」は、記憶装置１１５の各記憶部の最大転送レートのうち何％を使用しているかを示している。

【００４８】また、ネットワーク負荷情報テーブル４０３は、ネットワークインタフェース１１３を介してネットワーク１０１上に送出するデータが、どの計算機サイト（送出先サイト）に向けて送出され、どの程度の帯域幅（使用通信帯域幅［Ｍbps］）を使用しているか、また、受信しているデータがどの計算機サイト（送出元サイト）から送られて来たものであり、どの程度の帯域幅（使用通信帯域幅［Ｍbps］）を使用して受信しているかを示している。この「送出元サイト」の項目がデータの送出元の計算機サイトを示し、「送出先サイト」の項目がデータの送出先の計算機サイトを示している。すなわち、ＳｉｔｅＡは計算機サイトＡ１０５を、ＳｉｔｅＢは計算機サイトＢ１０６を、ＳｉｔｅＣは計算機サイトＣ１０７を示している。ここで、送信元サイト及び送信先サイトを総称してネットワークリンク（以下、単に「リンク」ともいう）という。また、「使用通信帯域幅」の項目が、送出元サイトと送出先サイトとの間で使用されている通信帯域幅［Ｍbps］を示している。

【００４９】すなわち、分散ファイル管理部１１２が分散ファイルＡの作成要求を受け取った場合、分散ファイル管理部１１２は、例えば図４に示すように、状態管理部１１４から獲得した記憶装置負荷情報テーブル４０２の情報に基づいて、記憶装置識別子ＤｉｓｋＩＤ１の記憶部の負荷が２０［％］で、残容量が１０［Ｍbytes］であるという情報を得る（ステップ３０１）。

【００５０】次に、分散ファイル管理部１１２は、状態管理部１１４から得られる記憶装置負荷情報テーブル４０２の情報に基づいて、サーバ計算機Ａ１０５に接続されている記憶装置１１５の中から、残容量が十分に残っており、且つ負荷が所定の閾値より低い記憶部を選択して、該記憶部に対して部分ファイルを順に割り当てていく。ここで、閾値としては、例えば、８０［％］などを使用するとよい。但し、この閾値は、記憶装置１１５の構成などに応じて適宜決定することができる。また、部分ファイルのサイズは、固定長が望ましく、全てのサーバ計算機で同一のサイズにするとよい。このとき、全ての部分ファイルＡ１～Ａ３を記憶装置１１５の記憶部に

割り当てられた場合にはステップ３０４の処理を行う。一方、全ての部分ファイルＡ１～Ａ３を記憶装置１１５の記憶部に割り当てられなかった場合には、ステップ３０３の処理を行う（ステップ３０２）。

【００５１】本実施の形態（図１）の場合では、全ての部分ファイルＡ１～Ａ３を記憶装置１１５の記憶部に割り当てられなかった（ステップ３０２）ので、分散ファイルＡの部分ファイルＡ１（１２６ａ）をサーバ計算機Ａ１０５に割り当て、残りの部分ファイルＡ２、Ａ３を他のサーバ計算機Ｃ１０７に割り当てている。

【００５２】分散ファイル管理部１１２は、ステップ３０２で割り当てられなかった部分ファイルＡ２、Ａ３を他のサーバ計算機に割り当てるために、他のサーバ計算機に対して内部ネットワーク１３１及びネットワーク１０１を介して、部分ファイルＡ２、Ａ３の作成を行なえるかどうかの問い合わせを行なう。問い合わせを受けた他のサーバ計算機では、自己の状態管理部の負荷情報テーブル４０１を調べ、部分ファイルＡ２、Ａ３の作成が可能かどうかを返答する（ステップ３０３）。この部分ファイルの作成の問い合わせや返答の信号のやり取りは、各サーバ計算機の分散ファイル管理部がネットワークインタフェース及びネットワークを介して行うことになる。

【００５３】本実施の形態（図１）の場合では、分散ファイル管理部１１２が、ネットワークインタフェース１１３及びネットワーク１０１を介して、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）の作成を行なえるかどうかを計算機サイトＣ１０４のサーバ計算機Ｃ１０７の分散ファイル管理部１２２に問い合わせる。サーバ計算機Ｃ１０７の分散ファイル管理部１２２は、状態管理部１２４から得られる負荷情報テーブル４０１に基づいて、上述のステップ３０２と同様な判断を行ない、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）の作成を行えるかどうかの返答をサーバ計算機Ａ１０５に行なう（ステップ３０３）。図１の場合、サーバ計算機Ｃ１０７の記憶装置１２５には、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）を作成することができる。

【００５４】次に、サーバ計算機Ａの分散ファイル管理部１１２は、サーバ計算機Ａの記憶装置１１５に割り当てられた部分ファイルＡ１と、他のサーバ計算機Ｃの記憶装置１２５に割り当てられた部分ファイルＡ２、Ａ３とを管理するための情報を登録する（ステップ３０４）。

【００５５】図５は、分散ファイル管理テーブル５０１を示す図である。また、図６は、部分ファイル管理テーブル６０１を示す図である。図５において、分散ファイル管理テーブル５０１は、分散ファイルを識別するための「分散ファイル識別子」と、分散ファイルを構成する部分ファイルを識別するための「部分ファイル識別子リスト」の項目で構成されている。また、図６において、部分ファイル管理テーブル６０１は、部分ファイルを識別するための「部分ファイル識別子」と、部分ファイルの所在地を示す「所在地」の項目で構成されている。ここで、図６に示した「部分ファイル識別子」は、図５で示した「部分ファイル識別子リスト」を構成する「部分ファイル識別子」に対応している。

【００５６】本実施の形態（図１）の場合では、例えば、分散ファイルＡについて見ると、図５において、分散ファイルＡが、部分ファイルＡ１（１２６ａ）、部分ファイルＡ２（１２６ｃ）、部分ファイルＡ３（１２６ｄ）から構成されることを表している。また、図６において、部分ファイルＡ１（１２６ａ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＡ／ｓｅｒｖｅｒＡ／ＤｉｓｋＩＤ１／（計算機サイトＡ１０２のサーバ計算機Ａ１０５の記憶装置識別子ＤｉｓｋＩＤ１）」であり、部分ファイルＡ２（１２６ｃ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／（計算機サイトＣ１０４のサーバ計算機Ｃ１０７の記憶装置識別子ＤｉｓｋＩＤ２）」であり、部分ファイルＡ３（１２６ｄ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／（計算機サイトＣ１０４のサーバＣ１０７の記憶装置ＤｉｓｋＩＤ２）」であることを表している。

【００５７】次に、分散ファイル管理部１１２は、サーバ計算機Ａの記憶装置１１５に部分ファイルＡ１を作成する場合、部分ファイル管理部１１１を介して、クライアント計算機１（１０８－１）からのデータを記憶装置１１５に書き込み、分散ファイルＡの部分ファイルＡ１の作成を行なう。また、他のサーバ計算機Ｃ１０７に部分ファイルＡ２、Ａ３を記録する場合、分散ファイル管理部１１２は、記録を行なうサーバ計算機Ｃ１０７の分散ファイル管理部１２２に部分ファイルＡ２、Ａ３の記録を依頼し、それと同時に、分散ファイルＡの作成要求を行なったクライアント計算機１（１０８－１）に指示をして、クライアント計算機１（１０８－１）から記録を行なうサーバ計算機Ｃ１０７に、直接データを送信するようにする。依頼を受けたサーバ計算機Ｃ１０７では、分散ファイル管理部１２２が部分ファイル管理テーブル６０１へ部分ファイルＡ２、Ａ３の登録を行なう。このようにして、他のサーバ計算機Ｃ１０７上に、部分ファイルＡ２、３が作成される（ステップ３０５）。

【００５８】以上のようにして、分散ファイルＡの部分ファイルＡ１（１２６ａ）の作成は、クライアント計算機１（１０８－１）からのデータを、サーバ計算機Ａ１０５の記憶装置１１５に書き込むことによって行なわれる。また、分散ファイルＡの部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）の作成は、クライアント計算機１（１０８－１）からのデータを、直接サーバ計算機Ｃ１０７に送り、サーバ計算機Ｃ１０７の

記憶装置１２５に書き込むことによって行なわれる。

【００５９】また、上述のステップ３０３において、他の全てのサーバ計算機で、部分ファイルの作成が不可能な場合には、分散ファイル管理部１１２は、分散ファイルの作成要求を行なったクライアント計算機１（１０８－１）に対して、分散ファイルの作成に失敗したことを通知する（ステップ３０６）。

【００６０】一方、ステップ３０２で、分散ファイルの全ての部分ファイルの作成を、自己のサーバ計算機の記憶装置にできる場合、分散ファイル管理部１１２は、記憶装置１１５に割り当てられた分散ファイルを管理するための情報を、図５に示した分散ファイル管理テーブル５０１と図６に示した部分ファイル管理テーブル６０１に登録する（ステップ３０７）。

【００６１】次に、分散ファイル管理部１１２は、部分ファイル管理部１１１を介して、クライアント計算機１（１０８－１）からのデータを記憶装置１１５に書き込み、全ての部分ファイルの作成を行なう（ステップ３０８）。

【００６２】以上のように、本発明の分散ファイル管理システムによれば、分散ファイルの作成を、各サーバ計算機の負荷を考慮して行うため、適切に負荷分散ができるようになる。

【００６３】次に、クライアント計算機群Ａ１０８内のクライアント計算機１（１０８－１）からサーバ計算機Ａ１０５に対して分散ファイル参照／更新要求が発行された場合について説明する。

【００６４】まず、クライアント計算機１（１０８－１）から発行された分散ファイルＡに対する参照／更新要求は、サーバ計算機Ａ１０５において、ネットワークインタフェース１１３を介して、分散ファイル管理部１１２によって受け取られる。

【００６５】図７は、分散ファイルの参照／更新要求を受け取った場合の分散ファイル管理部の動作アルゴリズムを示すフローチャートである。以下、図７を用いて、分散ファイル管理部１１２の詳細な動作を説明する。また、以下においては、分散ファイルＡに対する参照／更新要求の処理を例にした具体的な動作についても説明する。

【００６６】まず、分散ファイル管理部１１２は、クライアント計算機１（１０８－１）からの分散ファイルの参照／更新要求に基づいて、分散ファイル管理テーブル５０１と部分ファイル管理テーブル６０１から、参照／更新する部分ファイルを特定し、その部分ファイルの所在地を求める（ステップ７０１）。

【００６７】ここで、クライアント計算機１（１０８－１）からの分散ファイルＡに対する参照／更新要求の場合、分散ファイル管理部１１２は、分散ファイル管理テーブル５０１に基づいて、分散ファイルＡが、部分ファイルＡ１（１２６ａ）、部分ファイルＡ２（１２６ｃ）、部分ファイルＡ３（１２６ｄ）によって構成されていることが解る。また、部分ファイル管理テーブル６０１に基づいて、部分ファイルＡ１（１２６ａ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＡ／ｓｅｒｖｅｒＡ／ＤｉｓｋＩＤ１／（計算機サイトＡ１０２のサーバ計算機Ａ１０５の記憶装置識別子ＤｉｓｋＩＤ１）」であり、部分ファイルＡ２（１２６ｃ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／（計算機サイトＣ１０４のサーバ計算機Ｃ１０７の記憶装置識別子ＤｉｓｋＩＤ２）」であり、部分ファイルＡ３（１２６ｄ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／（計算機サイトＣ１０４のサーバＣ１０７の記憶装置ＤｉｓｋＩＤ２）」であることが解る。

【００６８】分散ファイル管理部１１２は、ステップ７０１で得られた部分ファイルの所在地から、参照／更新を行なう全ての部分ファイルが自己のサーバ計算機Ａ１０５の記憶装置１１５に存在するかどうか、あるいは一部または全部の部分ファイルが他のサーバ計算機に存在するかどうかの判定を行なう（ステップ７０２）。

【００６９】ここで、分散ファイルＡの場合、部分ファイルＡ１（１２６ａ）は、記憶装置１１５に存在し、部分ファィルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）は、計算機サイトＣ１０４のサーバ計算機Ｃ１０７の記憶装置１２５に存在することがわかる。

【００７０】次に、全ての部分ファイルが自己のサーバ計算機Ａ１０５の記憶装置１１５に存在しない場合（ステップ７０２）、ステップ７０１で得られた部分ファイルの所在地に基づいて、参照／更新を行なう部分ファイルが記録されている他のサーバ計算機に部分ファイルの存在を確かめる（ステップ７０３）。

【００７１】ここで、分散ファイルＡの場合には、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）の存在を、計算機サイトＣ１０４のサーバ計算機Ｃ１０７の分散ファイル管理部１２２に確認する。

【００７２】ステップ７０３で部分ファイルの存在が確認されたら、分散ファイル管理部１１２は、参照／更新を行なう部分ファイルが、自己のサーバ計算機Ａ１０５の記憶装置１１５に存在する場合には、クライアント計算機１（１０８－１）からの参照／更新要求に基づいて、部分ファイル管理部１１１を介して記憶装置１１５に存在する部分ファイルの読み出し（参照）や部分ファイルへの書き込み（更新）を行なう。また、参照／更新を行なう部分ファイルが、他のサーバ計算機の記憶装置に存在する場合、分散ファイル管理部１１２は、参照／更新を行なう部分ファイルを保持するサーバ計算機に該当部分ファイルの参照／更新を要求する。これと同時に、分散ファイル管理部１１２は、参照／更新要求を行なつたクライアント計算機１（１０８－１）が、参照／更新を行なう部分ファイルを保持するサーバ計算機に参

照／更新要求を直接行なうように指示する（ステップ７０４）。

【００７３】ここで、分散ファイルＡの場合、部分ファイルＡ１（１２６ａ）が計算機サイトＡ１０２のサーバ計算機Ａ１０５に、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）が計算機サイトＣ１０４のサーバ計算機Ｃ１０７に存在している。部分ファイルＡ１（１２６ａ）に対する参照／更新要求は、分散ファイル管理部１１２が、部分ファイル管理部１１１を介して、記憶装置１１５に対して参照／更新処理を行なう。一方、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）に対する参照／更新要求は、参照／更新要求を行ったクライアント計算機１（１０８－１）とサーバ計算機Ｃ１０７との間で、直接行われることになる。

【００７４】また、ステップ７０３で部分ファイルの存在が確認されなかった場合、分散ファイル管理部１１２は、分散ファイルの参照／変更要求を行ったクライアント計算機１（１０８－１）に、分散ファイルの参照／更新が失敗したことを通知する（ステップ７０５）。

【００７５】一方、ステップ７０２で、全ての部分ファイルが自己のサーバ計算機Ａ１０５の記憶装置１１５に存在する場合、分散ファイル管理部１１２は、クライアント計算機１（１０８－１）からの参照／更新要求に基づいて、部分ファイル管理部１１１を介して記憶装置１１５に存在する部分ファイルの読み出し（参照）や部分ファイルへの書き込み（更新）を行なう（ステップ７０６）。

【００７６】以上のように、上述した実施の形態によれば、クライアント計算機からサーバ計算機への要求が分散ファイルの作成の場合には、分散ファイルを複数の部分ファイルに分割し、サーバ計算機の負荷情報に基づいて各々の部分ファイルを作成するサーバ計算機を部分ファイル毎に決定して、分散ファイルの作成処理を行っている。また、クライアント計算機からの要求が分散ファイルの参照／更新の場合には、分散ファイルを構成する部分ファイルが存在するサーバ計算機を特定し、１または複数のサーバ計算機上に分散して配置されている部分ファイルをクライアント計算機から１つの分散ファイルとして扱うようにする。このようにして、クライアント計算機からサーバ計算機への分散ファイルの作成／参照／変更要求の際に、特定のサーバ計算機への負荷の集中をなくすことができる。

【００７７】（実施の形態２）図８は、本発明における分散ファイル管理システムの第２の実施の形態の一例を示す構成図である。この図８においては、図１と同様の構成には同一の符号を付している。図８に示した分散ファイル管理システムは、パーソナルコンピュータやワークステーションなどのサーバ計算機及びパーソナルコンピュータやワークステーションなどの複数のクライアント計算機から成るクライアント計算機群を備えた複数の計算機サイトＡ８０２、計算機サイトＢ８０３、及び計算機サイトＣ８０４と、計算機サイトＡ８０２、計算機サイトＢ８０３、及び計算機サイトＣ８０４を相互に接続するローカルエリアネットワークやワイドエリアネットワークなどのネットワーク１０１とを備えている。

【００７８】ここで、計算機サイトＡ８０２は、パーソナルコンピュータやワークステーションなどの複数のサーバ計算機（図８においては、「サーバ計算機Ａ８０５」のみ示す）と、パーソナルコンピュータやワークステーションなどのクライアント計算機１～ｎ（１０８－１～１０８－ｎ）から成るクライアント計算機群Ａ１０８とを備えている。この計算機サイトＡ８０２は、複数のサーバ計算機（図８においては、「サーバ計算機Ａ８０５」のみ示す）とクライアント計算機群Ａ１０８とをイーサネットなどの内部ネットワーク１３１で接続しており、例えば、インターネットドメインになっている。

【００７９】また、計算機サイトＡ８０２と同様に、計算機サイトＢ８０３は、複数のサーバ計算機（図８においては、「サーバ計算機Ｂ８０６」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｂ１０９とを備え、計算機サイトＣ８０４は、複数のサーバ計算機（図８においては、「サーバ計算機Ｃ８０７」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｃ１１０とを備えている。さらに、これらの計算機サイトＢ８０３及び計算機サイトＣ８０４は、計算機サイトＡ８０２と同様に、複数のサーバ計算機（図８においては、「サーバ計算機Ｂ８０６」及び「サーバ計算機Ｃ８０７」のみ示す）と、クライアント計算機群Ｂ１０９及びクライアント計算機群Ｃ１１０とを、それぞれ内部ネットワーク１３２及び内部ネットワーク１３３で接続しており、例えば、インターネットドメインになっている。

【００８０】サーバ計算機Ａ８０５は、分散ファイルの部分ファイルを記録するハードディスクなどの記憶装置１１５と、イーサネットなどの内部ネットワーク１３１へ接続するためのネットワークインタフェース１１３と、部分ファイルを記録している記憶装置１１５への書き込みや読み出しを制御する部分ファイル管理部１１１と、記憶装置１１５に対する負荷や記憶装置１１５の残り容量、及びネットワークインタフェース１１３に対する負荷を監視し、これらの負荷や容量に関する情報を保持する状態管理部８１４と、部分ファイル管理部１１１、状態管理部８１４、及びネットワークインタフェース１１３に接続された分散ファイル管理部１１２とによって構成されている。

【００８１】この状態管理部８１４は、他のサーバ計算機へ負荷情報を通知し、また、他のサーバ計算機から通知された外部負荷情報を保持する外部状態管理部８１１を備えている。

【００８２】サーバ計算機Ｂ８０６及びサーバ計算機Ｃ８０７は、サーバ計算機Ａ８０５と同様の構成になっている。すなわち、サーバ計算機Ｂ８０６は、記憶装置１２０と、ネットワークインタフェース１１８と、部分ファイル管理部１１６と、外部状態管理部８１２を備えた状態管理部８１９と、分散ファイル管理部１１７とによって構成されている。また、サーバ計算機Ｃ８０７は、記憶装置１２５と、ネットワークインタフェース１２３と、部分ファイル管理部１２１と、外部状態管理部８１３を備えた状態管理部８２４と、分散ファイル管理部１２２とによって構成されている。

【００８３】ここで、図８に示した分散ファイル管理システムと図１に示した分散ファイル管理システムとの相違点は、図８に示した状態管理部８１４、８１９、８２４が、他のサーバ計算機へ負荷情報を通知し、他のサーバ計算機から通知された外部負荷情報を保持する外部状態管理部８１１、８１２、８１３を備えている点である。

【００８４】図９は、外部状態管理部８１１、８１２、８１３で管理されている外部負荷情報テーブル９０１の一例を示す。図９において、外部負荷情報テーブル９０１は、サーバ計算機の所在地を示す「サーバ計算機所在地」と、サーバ計算機所在地で示されるサーバ計算機の記憶装置の負荷情報を示す「記憶装置負荷情報」の項目で構成されている。また、「記憶装置負荷情報」は、記憶装置を識別するための「記憶装置識別子」と、記憶装置の負荷を示す「負荷」と、記憶装置の残容量を示す「残容量」の項目で構成されている。

【００８５】外部状態管理部８１１、８１２、８１３は、外部負荷情報テーブル９０１からの外部負荷情報を他のサーバ計算機へ通知するために、以下のような動作を行なう。

【００８６】まず、外部状態管理部８１１、８１２、８１３は、定期的または所定のタイミングで状態管理部８１４、８１９、８２４に対して記憶装置負荷情報テーブル４０２に示される情報を問い合わせるか、または、状態管理部８１４、８１９、８２４から記憶装置負荷情報テーブル４０２に示される情報の状態変化を通知してもらうことにより、状態管理部８１４、８１９、８２４で管理されている記憶装置負荷情報テーブル４０２の情報を得る。

【００８７】次に、外部状態管理部８１１、８１２、８１３は、それぞれのネットワークインタフェース１１３、１１８、１２３を介して、各サーバ計算機に、記憶装置負荷情報を通知する。通知を受けたサーバ計算機では、それぞれネットワークインタフェース１１３、１１８、１２３を介して、外部状態管理部８１１、８１２、８１３が、記憶装置負荷情報を受け取り、それぞれの外部負荷情報テーブル９０１にこの情報を記録していく。

【００８８】以上のように構成された分散ファイル管理システムについて、クライアント計算機群Ａ１０８内のクライアント計算機１（１０８－１）からサーバ計算機Ａ８０５に対して分散ファイルＡの作成要求が発行された場合を例として説明する。

【００８９】まず、クライアント計算機１（１０８－１）から発行された分散ファイルＡの作成要求は、内部ネットワーク１３１及びサーバ計算機Ａ８０５のネットワークインタフェース１１３を介して、分散ファイル管理部１１２によって受け取られる。

【００９０】図１０は、分散ファイルの作成要求を受け取ったときの分散ファイル管理部の動作アルゴリズムを示すフローチャートである。以下、図１０を用いて、分散ファイル管理部の詳細な動作を説明する。本実施の形態では、第１の実施の形態で説明した図３に示すステップ３０２とステップ３０３の処理を統合して、１つのステップ１００２で処理することができる。

【００９１】まず、分散ファイル管理部１１２は、状態管理部８１４の管理している負荷情報テーブル４０１と、外部状態管理部８１１の管理している外部負荷情報テーブル９０１から各情報を獲得する（ステップ１００１）。

【００９２】この状態管理部１１４では、図４に示したような負荷情報テーブル４０１を管理している。図４において、負荷情報テーブル４０１は、記憶装置負荷情報テーブル４０２と、ネットワーク負荷情報テーブル４０３とによって構成されている。

【００９３】この記憶装置負荷情報テーブル４０２は、サーバ計算機Ａ１０５に接続されている記憶装置１１５の複数の記憶部を識別するための「記憶装置識別子」と、各記憶部の負荷情報［％］を示す「負荷」と、各記憶部の残容量［Mbytes］を示す「残容量」の項目で構成されている。ここで、記憶装置１１５の各記憶部の「負荷」は、記憶装置１１５の各記憶部の最大転送レートのうち何％を使用しているかを示している。

【００９４】また、ネットワーク負荷情報テーブル４０３は、ネットワークインタフェース１１３を介してネットワーク１０１上に送出するデータが、どの計算機サイト（送出先サイト）に向けて送出され、どの程度の帯域幅（使用通信帯域幅［Mbps］）を使用しているか、また、受信しているデータがどの計算機サイト（送出元サイト）から送られて来たものであり、どの程度の帯域幅（使用通信帯域幅［Mbps］）を使用して受信しているかを示している。この「送出元サイト」の項目がデータの送出元の計算機サイトを示し、「送出先サイト」の項目がデータの送出先の計算機サイトを示している。すなわち、ＳｉｔｅＡは計算機サイトＡ１０５を、ＳｉｔｅＢは計算機サイトＢ１０６を、ＳｉｔｅＣは計算機サイトＣ１０７を示している。ここで、送信元サイト及び送信先サイトを総称してネットワークリンク（以下、単に「リンク」ともいう）という。また、「使用通信帯域

幅」の項目が、送出元サイトと送出先サイトとの間で使用されている通信帯域幅［Mbps］を示している。

【００９５】また、外部状態管理部８１１は、図９に示したような外部負荷情報テーブル９０１を管理している。

【００９６】ここで、分散ファイルＡを作成する場合、分散ファイル管理部１１２は、状態管理部８１４の負荷情報テーブル４０１から、「記憶装置識別子」がＤｉｓｋＩＤ１で示される記憶装置１１５の記憶部の「負荷」が２０［％］で、「残容量」が１０［Mbytes］であるという情報を得ることができる。また、分散ファイル管理部１１２は、外部状態管理部８１４の外部負荷情報テーブル９０１から、計算機サイトＢ８０３のサーバ計算機Ｂ８０６の「記憶装置識別子」がＤｉｓｋＩＤ１で示される記憶装置１２０の記憶部の「負荷」が４９［％］で、「残容量」が１０００［Mbytes］であり、また、計算機サイトＣ８０４のサーバ計算機Ｃ８０７の「記憶装置識別子」がＤｉｓｋＩＤ１で示される記憶装置１２５の記憶部の「負荷」が３０［％］で、「残容量」が３０００［Mbytes］であるという情報を得ることができる。

【００９７】次に、分散ファイル管理部１１２は、まず、状態管理部８１４から得られる記憶装置負荷情報テーブル４０２に基づいて、サーバ計算機Ａ８０５に接続されている記憶装置１１５の各記憶部の中から、「残容量」が所定の容量以上で、且つ「負荷」が所定の閾値より低いという条件を満たす記憶部を選択し、分散ファイルを分割した部分ファイルを、当該条件を満たす各記憶部に順に割り当てていく。全ての部分ファイルを記憶装置１１５に割り当てられない場合には、分散ファイル管理部１１２は、外部状態管理部８１１から得られる外部負荷情報テーブル９０１の情報に基づいて、記憶部の「残容量」が所定の容量以上で、且つ「負荷」が所定の閾値より低い記憶部の存在する記憶装置を持つ他のサーバ計算機を選択する。そして、まだ割り当てられていない部分ファイルを、当該他のサーバ計算機の記憶装置の記憶部に順に割り当て、この割当てを当該他のサーバ計算機に通知する。そして、全ての部分ファイルがサーバ計算機の各記憶装置の記憶部に割り当てて作成できたか否かを判断する（ステップ１００２）。

【００９８】ここで、分散ファイルＡを作成する場合、分散ファイル管理部１１２は、状態管理部８１４の記憶装置負荷情報テーブル４０１に基づいて、部分ファイルＡ１（１２６ａ）をサーバ計算機Ａ８０５の記憶装置１１５の所定の記憶部に割り当てる。また、分散ファイル管理部１１２は、外部状態管理部８１１の外部負荷情報テーブル９０１の記憶装置負荷情報に基づいて、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）をサーバ計算機Ｃ８０７の記憶装置１２５の各記憶部に割り当てる。

【００９９】このように、サーバ計算機Ａ８０５は、外部状態管理部８１１から他のサーバ計算機の負荷や残容量の情報を得ることにより、他のサーバ計算機に部分ファイルの作成が可能かどうかの問い合わせをすることなく、他のサーバ計算機の負荷や残容量を考慮しながら部分ファイルを配置（記憶）するサーバ計算機を決めることができる。

【０１００】次に、分散ファイル管理部１１２は、分散ファイルの管理情報を、図５に示したような分散ファイル管理テーブル５０１と、図６に示したような部分ファイル管理テーブル６０１に登録する（ステップ１００３）。図５において、分散ファイル管理テーブル５０１は、分散ファイルを識別するための「分散ファイル識別子」と、分散ファイルを構成する部分ファイルを識別するための「部分ファイル識別子リスト」の項目で構成されている。また、図６において、部分ファイル管理テーブル６０１は、部分ファイルを識別するための「部分ファイル識別子」と、部分ファイルの所在地を示す「所在地」の項目で構成されている。ここで、図６に示した「部分ファイル識別子」は、図５で示した「部分ファイル識別子リスト」を構成する「部分ファイル識別子」に対応している。

【０１０１】ここで、分散ファイルＡの場合、図５において、分散ファイルＡが、部分ファイルＡ１（１２６ａ）、部分ファイルＡ２（１２６ｃ）、部分ファイルＡ３（１２６ｄ）から構成されることを表している。また、図６において、部分ファイルＡ１（１２６ａ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＡ／ｓｅｒｖｅｒＡ／ＤｉｓｋＩＤ１／（計算機サイトＡ１０２のサーバ計算機Ａ１０５の記憶装置識別子ＤｉｓｋＩＤ１）」であり、部分ファイルＡ２（１２６ｃ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／（計算機サイトＣ１０４のサーバ計算機Ｃ１０７の記憶装置識別子ＤｉｓｋＩＤ２）」であり、部分ファイルＡ３（１２６ｄ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／（計算機サイトＣ１０４のサーバＣ１０７の記憶装置ＤｉｓｋＩＤ２）」であることを表している。

【０１０２】次に、分散ファイル管理部１１２は、記憶装置１１５の記憶部に部分ファイルを記録する場合には、部分ファイル管理部１１１を介して、クライアント計算機１（１０８－１）からのデータを記憶装置１１５の所定の記憶部に書き込む。また、他のサーバ計算機の記憶装置に部分ファイルを記録する場合には、分散ファイル管理部１１２は、記録を行なうサーバ計算機の分散ファイル管理部に部分ファイルの記録を依頼する。これと同時に、分散ファイル管理部１１２は、分散ファイルの作成要求を行なったクライアント計算機１（１０８－１）に指示して、記録を行なうサーバ計算機に直接データを送信するように指示する。分散ファイル管理部１１２から依頼を受けたサーバ計算機では、クライアント計

算機１（１０８－１）から部分ファイルのデータを受け取って、記憶装置の所定の記憶部に記録する。また、該サーバ計算機の分散ファイル管理部は、部分ファイル管理部の部分ファイル管理テーブル６０１へ部分ファイルの情報の登録を行なう。このようにして、他のサーバ計算機上に、部分ファイルが作成される（ステップ１００４）。

【０１０３】分散ファイルＡの場合、部分ファイルＡ１（１２６ａ）の作成は、クライアント計算機１（１０８－１）からのデータを記憶装置１１５の所定の記憶部に書き込むことによって行なわれる。部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）の作成は、クライアント計算機１（１０８－１）から所定のデータを直接サーバ計算機Ｃ８０７に送り、サーバ計算機Ｃ８０７の記憶装置１２５の所定の記憶部にそれぞれ書き込むことによって行なわれる。

【０１０４】一方、ステップ１００２で、どのサーバ計算機にも部分ファイルを作成できない場合には、分散ファイル管理部１１２は、分散ファイル作成の要求を行なったクライアント計算機１（１０８－１）に対して、分散ファイルの作成に失敗したことを通知する（ステップ１００５）。

【０１０５】以上、分散ファイルの作成について説明したが、クライアント計算機からサーバ計算機に対して分散ファイルの参照／更新要求が発行された場合については、第１の実施の形態の場合（図７）と同様である。

【０１０６】以上のように、本発明の第２の実施の形態においては、状態管理部８１４、８１９、８２４が、他のサーバ計算機へ負荷情報を通知し、また、他のサーバ計算機から通知された外部負荷情報を保持する外部状態管理部８１１、８１２、８１３を備えることにより、分散ファイル管理部１１２、１１７、１２２は、他のサーバ計算機の負荷情報に基づいて、分散ファイルの部分ファイルを配置するサーバ計算機を決定することができ、特定のサーバ計算機への負荷の集中を回避することができる。

【０１０７】なお、外部状態管理部８１１、８１２、８１３から各サーバ計算機に記憶装置負荷情報を通知する際には、ユニキャストやマルチキャストを用いるとよい。特に、マルチキャストを用いた場合、全サーバ計算機に記憶装置負荷情報を一斉に通知することができ、通知のための通信量を減らすことができる。

【０１０８】また、通知するサーバ計算機をあらかじめ複数のグループに分けておき、各グループに属するサーバ計算機の各々にユニキャストで通知することもでき、また、各グループに対してマルチキャストで通知することもできる。このようにして、通知のための通信量を減らすことができる。

【０１０９】さらに、通知するサーバ計算機を隣接するサーバ計算機、すなわち、ネットワークで直接に接続されているサーバ計算機に限定して、ユニキャストあるいはマルチキャストによって通知することもできる。これにより、通知のための通信量を減らすことができる。

【０１１０】（実施の形態３）図１１は、本発明における分散ファイル管理システムの第３の実施の形態の一例を示す構成図である。この図１１においては、図８と同様の構成には同一の符号を付している。図１１に示した分散ファイル管理システムは、パーソナルコンピュータやワークステーションなどのサーバ計算機及びパーソナルコンピュータやワークステーションなどの複数のクライアント計算機から成るクライアント計算機群を備えた複数の計算機サイトＡ１１０２、計算機サイトＢ１１０３、及び計算機サイトＣ１１０４と、計算機サイトＡ１１０２、計算機サイトＢ１１０３、及び計算機サイトＣ１１０４を相互に接続するローカルエリアネットワークやワイドエリアネットワークなどのネットワーク１０１とを備えている。

【０１１１】ここで、計算機サイトＡ１１０２は、パーソナルコンピュータやワークステーションなどの複数のサーバ計算機（図１１においては、「サーバ計算機Ａ１１０５」のみ示す）と、パーソナルコンピュータやワークステーションなどのクライアント計算機１～ｎ（１０８－１～１０８－ｎ）から成るクライアント計算機群Ａ１０８とを備えている。この計算機サイトＡ１１０２は、複数のサーバ計算機（図１１においては、「サーバ計算機Ａ１１０５」のみ示す）とクライアント計算機群Ａ１０８とをイーサネットなどの内部ネットワーク１３１で接続しており、例えば、インターネットドメインになっている。

【０１１２】また、計算機サイトＡ１１０２と同様に、計算機サイトＢ１１０３は、複数のサーバ計算機（図１１においては、「サーバ計算機Ｂ１１０６」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｂ１０９とを備え、計算機サイトＣ１１０４は、複数のサーバ計算機（図１１においては、「サーバ計算機Ｃ１１０７」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｃ１１０とを備えている。さらに、これらの計算機サイトＢ１１０３及び計算機サイトＣ１１０４は、計算機サイトＡ１１０２と同様に、複数のサーバ計算機（図１１においては、「サーバ計算機Ｂ１１０６」及び「サーバ計算機Ｃ１１０７」のみ示す）と、クライアント計算機群Ｂ１０９及びクライアント計算機群Ｃ１１０とを、それぞれ内部ネットワーク１３２及び内部ネットワーク１３３で接続しており、例えば、インターネットドメインになっている。

【０１１３】サーバ計算機Ａ１１０５は、分散ファイルの部分ファイルを記録するハードディスクなどの記憶装置１１５と、イーサネットなどの内部ネットワーク１３１へ接続するためのネットワークインタフェース１１３と、部分ファイルを記録している記憶装置１１５への書

き込みや読み出しを制御する部分ファイル管理部１１１と、記憶装置１１５に対する負荷や記憶装置１１５の残り容量、及びネットワークインタフェース１１３に対する負荷を監視し、これらの負荷や容量に関する情報を保持する状態管理部８１４と、部分ファイル管理部１１１、状態管理部８１４、及びネットワークインタフェース１１３に接続された分散ファイル管理部１１１２とによって構成されている。

【０１１４】この状態管理部８１４は、他のサーバ計算機へ負荷情報を通知し、また、他のサーバ計算機から通知された外部負荷情報を保持する外部状態管理部８１１を備えている。

【０１１５】また、分散ファイル管理部１１１２は、部分ファイル毎のアクセス情報と負荷情報テーブル４０１及び外部負荷情報テーブル９０１の情報とに基づいて、移動させる部分ファイルを決定し、他のサーバ計算機へ部分ファイルを移動させる分散ファイル移動部１１３１を備えている。

【０１１６】サーバ計算機Ｂ１１０６及びサーバ計算機Ｃ１１０７は、サーバ計算機Ａ１１０５と同様の構成になっている。すなわち、サーバ計算機Ｂ１１０６は、記憶装置１２０と、ネットワークインタフェース１１８と、部分ファイル管理部１１６と、外部状態管理部８１２を備えた状態管理部８１９と、分散ファイル移動部１１３２を備えた分散ファイル管理部１１１７とによって構成されている。また、サーバ計算機Ｃ１１０７は、記憶装置１２５と、ネットワークインタフェース１２３と、部分ファイル管理部１２１と、外部状態管理部８１３を備えた状態管理部８２４と、分散ファイル移動部１１３３を備えた分散ファイル管理部１１２２とによって構成されている。

【０１１７】ここで、図１１に示した分散ファイル管理システムと図８に示した分散ファイル管理システムとの相違点は、図１１に示した分散ファイル管理部１１１２、１１１７、１１２２が、部分ファイル毎のアクセス情報と負荷情報テーブル４０１及び外部負荷情報テーブル９０１の情報とから、移動させる部分ファイルを決定し、他のサーバ計算機へ部分ファイルを移動させる分散ファイル移動部１１３１、１１３２、１１３３を備えている点である。

【０１１８】図１２は、状態管理部８１４で管理されている部分ファイル毎のアクセス情報テーブル１２０１の一例を示す。図１２において、このアクセス情報テーブル１２０１は、部分ファイルを識別するための「部分ファイル識別子」と、「単位時間あたりのアクセス情報」の項目で構成されている。また、「単位時間あたりのアクセス情報」は、部分ファイルにアクセスしているクライアント計算機の存在するサイトの情報である「アクセス元サイト識別子」と、部分ファイルへのアクセス回数を示す「アクセス回数」の項目から構成されている。このアクセス情報テーブル１２０１は、状態管理部８１４によって、単位時間毎に更新され続ける。

【０１１９】図１３は、分散ファイル管理部１１１２で管理されている部分ファイル管理テーブル１３０１の一例を示す。図１３において、部分ファイル管理テーブル１３０１は、部分ファイルを識別するための「部分ファイル識別子」と、部分ファイルの所在地を示す「所在地」と、部分ファイルが最初に作成された所在地を示す「オリジナル所在地」の項目から構成されている。図１３において、部分ファイルが作成された段階では、「所在地」と「オリジナル所在地」の示す情報は同一であるが、部分ファイルが他のサーバ計算機に移動するのに応じて、「所在地」の情報は変化する。図１３で示した部分ファイル管理テーブル１３０１は、図６で示した部分ファイル管理テーブル６０１に、「オリジナル所在地」の項目を加えたものになっている。

【０１２０】以上のように構成された分散ファイル管理システムにおいて、分散ファイルＡ、分散ファイルＢ、及び分散ファイルＣが、上述で示したようにしてサーバ計算機Ａ１１０５で作成された後、各分散ファイルの部分ファイルを移動する処理について詳細に説明する。

【０１２１】図１４は、サーバ計算機Ａ１１０５によって作成された分散ファイルＡ、分散ファイルＢ、及び分散ファイルＣの部分ファイル管理テーブル１３０１の内容の一例を示している。図１４において、図１４（Ａ）は、サーバ計算機Ａ１１０５の部分ファイル管理テーブル１４０１を示し、（Ｂ）はサーバ計算機Ｂ１１０６の部分ファイル管理テーブル１４０２を示し、（Ｃ）はサーバ計算機Ｃ１１０７の部分ファイル管理テーブル１４０３を示している。図１４（Ａ）の部分ファイル管理テーブル１４０１には、部分ファイル識別子が、Ａ１、Ａ２、Ａ３、Ｂ１、Ｃ１、Ｃ２で示される各部分ファイルの所在地と、オリジナル所在地が示されている。部分ファイル管理テーブル１４０２には、部分ファイル識別子が、Ｃ１、Ｃ２で示される各部分ファイルの所在地と、オリジナル所在地が示されている。部分ファイル管理テーブル１４０３には、部分ファイル識別子が、Ａ２、Ａ３で示される各部分ファイルの所在地と、オリジナル所在地が示されている。ここで、図１４においては、各部分ファイルの移動前の状態を表している。このため、全ての部分ファイルにおいて、その所在地とオリジナル所在地が一致している。

【０１２２】図１４に示した状態での部分ファイルの移動の際の、サーバ計算機Ａ１１０５の分散ファイル移動部１１３１の動作アルゴリズムについて説明する。

【０１２３】図１５は、分散ファイル移動部１１３１の動作アルゴリズムを示す。まず、分散ファイル移動部１１３１は、状態管理部８１４が管理している負荷情報テーブル４０１（図４）の情報を、一定時間の間隔で監視する（ステップ１５０１）。

【０１２４】分散ファイル移動部１１３１は、ある記憶装置の「負荷」が予め設定されている所定の閾値（例えば、８０％などの値で、この値は、システムの構成などによって任意に決定する）を越えたことを検出すると（ステップ１５０１）、この検出された記憶装置に含まれている部分ファイルの「部分ファイル識別子」を、部分ファイル管理テーブル１３０１を参照して探す。探し出した「部分ファイル識別子」の「単位時間当たりのアクセス情報」を、アクセス情報テーブル１２０１から得る。ここで得られた「単位時間当たりのアクセス情報」の「アクセス回数」を各「部分ファイル識別子」毎に比較し、最も大きな「アクセス回数」になっている「部分ファイル識別子」を選択する（ステップ１５０２）。すなわち、移動元部分ファイルを選択する。例えば、ここで、移動元部分ファイルとして部分ファイルＡ１（１２６ａ）が選択されたとする。

【０１２５】次に、分散ファイル移動部１１３１は、外部負荷情報テーブル９０１（図９）に基づいて、「残容量」が十分で且つ「負荷」が所定の値より低い記憶装置を持つサーバ計算機を選択する。そして、分散ファイル移動部１１３１は、選択したサーバ計算機に対して部分ファイルが移動できるかどうかを確認し、移動可能なサーバ計算機を決定する（ステップ１５０２）。すなわち、移動先サーバ計算機を選択する。例えば、ここで、サーバ計算機Ｃ１１０７の記憶装置識別子ＤｉｓｋＩＤ２で示される記憶装置１２５の記憶部が選択されたとする。

【０１２６】次に、この移動元部分ファイルの選択と移動先サーバ計算機の選択によって得られた情報に基づいて、部分ファイルの移動を行なう（ステップ１５０３）。

【０１２７】上述の例では、ステップ１５０２で、移動元部分ファイルとして部分ファイルＡ１（１２６ａ）、移動先サーバ計算機としてサーバ計算機Ｃ１１０７が選択されているので、移動元のサーバ計算機Ａ１１０５の分散ファイル移動部１１３１は部分ファイル管理部１１１を介して、記憶装置１１５から部分ファイルＡ１（１２６ａ）を読み出す。この読み出された部分ファイルＡ１（１２６ａ）は、部分ファイルＡ１（１２６ａ）の「オリジナル所在地」（図１３）に関する情報と共に、ネットワークインタフェース１１３及び内部ネットワーク１３１を介してネットワーク１０１へ送出される。

【０１２８】一方、移動先のサーバ計算機Ｃ１１０７において、分散ファイル移動部１１３３は、移動元のサーバ計算機Ａ１１０５より送出された部分ファイルＡ１（１２６ａ）とその「オリジナル所在地」の情報を、ネットワーク１０１から内部ネットワーク１３３及びネットワークインタフェース１２３を介して受信する。部分ファイル管理部１２１は、この受信した部分ファイルＡ１（１２６ａ）を記憶装置１２５に書き込む。また、部分ファイルＡ１（１２６ａ）の「オリジナル所在地」を、部分ファイル管理部１２１の部分ファイル管理テーブル（図１４（Ｃ））に登録する。

【０１２９】その後、移動先のサーバ計算機Ｃ１１０７は、移動元のサーバ計算機Ａ１１０５と「オリジナル所在地」（図１３）に示されているサーバ計算機（この例の場合では、サーバ計算機Ａ１１０５、すなわち、移動元とオリジナルは同じサーバ計算機Ａ１１０５）に対して、部分ファイルＡ１（１２６ａ）の移動が完了したことを通知する。移動元のサーバ計算機Ａ１１０５と「オリジナル所在地」に示されるサーバ計算機では、部分ファイル管理テーブルに登録されている部分ファイルＡ１（１２６ａ）の情報を書き換える。

【０１３０】図１６は、上述の例のように部分ファイルＡ１（１２６ａ）が移動した後の、図１４に示した部分ファイル管理テーブル１４０１、１４０２、１４０３の状態を示す。図１６において、図１６（Ａ）はサーバ計算機Ａ１１０５の部分ファイル管理テーブル１６０１を示し、（Ｂ）はサーバ計算機Ｂ１１０６の部分ファイル管理テーブル１６０２を示し、（Ｃ）はサーバ計算機Ｃ１１０７の部分ファイル管理テーブル１６０３を示している。すなわち、図１６（Ａ）～（Ｃ）の各部分ファイル管理テーブル１６０１、１６０２、１６０３は、それぞれ図１４（Ａ）～（Ｃ）の各部分ファイル管理テーブル１４０１、１４０２、１４０３に対応している。ここで、部分ファイル管理テーブル１６０１には、「部分ファイル識別子」がＡ１、Ａ２、Ａ３、Ｂ１、Ｃ１、Ｃ２で示される各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。部分ファイル管理テーブル１６０２には、「部分ファイル識別子」がＣ１、Ｃ２で示される各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。部分ファイル管理テーブル１６０３には、「部分ファイル識別子」がＡ１、Ａ２、Ａ３で示される各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。ここで、図１６に示した部分ファイル管理テーブル１６０１、１６０２、１６０３の状態と図１４に示した部分ファイル管理テーブル１４０１、１４０２、１４０３の状態の差異は、部分ファイルＡ１（１２６ａ）を、サーバ計算機Ａ１１０５からサーバ計算機Ｃ１１０７へ移動させたことによるものである。すなわち、図１６と図１４の相違点は、図１４（Ａ）の部分ファイル管理テーブル１４０１において、部分ファイルＡ１（１２６ａ）の「所在地」が「ｆｉｌｅ：／／ｓｉｔｅＡ／ｓｅｒｖｅｒＡ／ＤｉｓｉＩＤ１／」であると登録されている情報が、図１６（Ａ）の部分ファイル管理テーブル１６０１においては、部分ファイルＡ１（１２６ａ）の「所在地」が「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／」と登録されている点と、図１６（Ｃ）の部分ファイル管理テーブル１６０３において部分ファイルＡ１の項目が追加

されている点である。

【０１３１】図１７は、さらに、この図１６に示した状態から、上述した処理と同様にして、サーバ計算機Ｂ１１０６の分散ファイル移動部１１３２が、部分ファイルＣ１（１２６ｅ）を、サーバ計算機Ｃ１１０７に移動させ、サーバ計算機Ｃ１１０７の分散ファイル移動部１１３３が、部分ファイルＡ１（１２６ａ）を、サーバ計算機Ｂ１１０６に移動させた場合の部分ファイル管理テーブルを示す。

【０１３２】図１７において、図１７（Ａ）はサーバ計算機Ａ１１０５の部分ファイル管理テーブル１７０１を示し、（Ｂ）はサーバ計算機Ｂ１１０６の部分ファイル管理テーブル１７０２を示し、（Ｃ）はサーバ計算機Ｃ１１０７の部分ファイル管理テーブル１７０３を示している。部分ファイル管理テーブル１７０１には、「部分ファイル識別子」がＡ１、Ａ２、Ａ３、Ｂ１、Ｃ１、Ｃ２で示される各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。部分ファイル管理テーブル１７０２には、「部分ファイル識別子」がＣ１、Ｃ２、Ａ１で示される各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。また、部分ファイル管理テーブル１７０３には、「部分ファイル識別子」がＡ２、Ａ３、Ｃ１で示される各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。

【０１３３】ここで、図１７に示した部分ファイル管理テーブル１７０１、１７０２、１７０３の状態と図１６に示した部分ファイル管理テーブル１６０１、１６０２、１６０３の状態の差異は、部分ファイルＡ１（１２６ａ）を、サーバ計算機Ｃ１１０７からサーバ計算機Ｂ１１０６へ移動させたことと、部分ファイルＣ１（１２６ｅ）をサーバ計算機Ｂ１１０６からサーバ計算機Ｃ１１０７へ移動させたによるものである。

【０１３４】すなわち、部分ファイルＡ１（１２６ａ）の移動に対応して、部分ファイル管理テーブル１７０２には、部分ファイルＡ１（１２６ａ）の項目が追加されている。また、部分ファイル管理テーブル１７０３においては、部分ファイルＡ１（１２６ａ）の項目（図１６参照）が削除されている。さらに、サーバ計算機Ｂ１１０６が部分ファイルＡ１（１２６ａ）の「オリジナル所在地」を参照して、サーバ計算機Ａ１１０５に移動を知らせ、この通知によって、「オリジナル所在地」で示されるサーバ計算機Ａ１１０５においては、部分ファイル管理テーブル１７０１に登録されている部分ファイルＡ１（１２６ａ）の「所在地」を「ｆｉｌｅ：／／ｓｉｔｅＢ／ｓｅｒｖｅｒＢ／ＤｉｓｋＩＤ２／」に変更している。

【０１３５】また、部分ファイルＣ１（１２６ｅ）の移動に対応して、部分ファイル管理テーブル１７０２に登録されている部分ファイルＣ１（１２６ｅ）の「所在地」が、「ｆｉｌｅ：／／ｓｉｔｅＢ／ｓｅｒｖｅｒＢ／ＤｉｓｋＩＤ３／」（図１６（Ｂ））から、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」に変更されている。また、部分ファイル管理テーブル１７０３においては、部分ファイルＣ１（１２６ｅ）の項目が追加されている。

【０１３６】次に、図１７に示す状態で、クライアント計算機１（１０８－１）が、分散ファイルＣを参照する際に、その参照内容が部分ファイルＣ１（１２６ｅ）に含まれる場合の動作を説明する。

【０１３７】（１）クライアント計算機１（１０８－１）は、分散ファイルＣを作成したサーバ計算機Ａ１１０５に対して、分散ファイルＣの参照を要求する。サーバ計算機Ａ１１０５では、分散ファイル管理テーブル１７０１を参照して、分散ファイルＣを構成する部分ファイルＣ１、Ｃ２のうちどの部分ファイルを参照しているのかを調べる。ここでは、部分ファイルＣ１（１２６ｅ）とする。部分ファイルＣ１（１２６ｅ）の「所在地」は、「ｆｉｌｅ：／／ｓｉｔｅＢ／ｓｅｒｖｅｒＢ／ＤｉｓｋＩＤ３／」なので、サーバ計算機Ａ１１０５は、サーバ計算機Ｂ１１０６に対して、部分ファイルＣ１（１２６ｅ）が存在するかどうかの確認を行なう。

【０１３８】（２）サーバ計算機Ｂ１１０６は、分散ファイル管理テーブル１７０２を調べて、部分ファイルＣ１（１２６ｅ）の「所在地」を調べる。部分ファイルＣ１（１２６ｅ）の「所在地」は、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」なので、サーバ計算機Ｂ１１０６は、サーバ計算機Ｃ１１０７に対して、部分ファイルＣ１（１２６ｅ）が存在するかどうかの確認を行なう。

【０１３９】（３）サーバ計算機Ｃ１１０７は、分散ファイル管理テーブル１７０３を調べて、部分ファイルＣ１（１２６ｅ）の「所在地」を調べる。部分ファイルＣ１（１２６ｅ）の「所在地」は、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」なので、部分ファイルＣ１（１２６ｅ）は、サーバ計算機Ｃ１１０７に存在することが解る。

【０１４０】（４）サーバ計算機Ｃ１１０７は、サーバ計算機Ｂ１１０６に、部分ファイルＣ１（１２６ｅ）は、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」に存在することを通知する。

【０１４１】（５）この通知を受けて、サーバ計算機Ｂ１１０６は、サーバ計算機Ａ１１０５に、部分ファイルＣ１（１２６ｅ）が「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」に存在することを通知する。

【０１４２】（６）サーバ計算機Ａ１１０５は、サーバ計算機Ｃ１１０７に部分ファイルＣ１（１２６ｅ）の参照を要求する。この要求と同時に、参照を要求したクライアント計算機１（１０８－１）に対して、サーバ計算機Ｃ１１０７に直接、部分ファイルＣ１（１２６ｅ）の

参照要求を行なうように指示する。また、サーバ計算機Ａ１１０５では、部分ファイルＣ１（１２６ｅ）の「所在地」を、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」へと書き換える。

【０１４３】以上のように、本実施の形態では、分散ファイル管理部１１１２、１１１７、１１２２が、部分ファイル毎のアクセス情報テーブル１２０１、負荷情報テーブル４０１及び外部負荷情報テーブル９０１の各情報に基づいて、移動させる部分ファイルを決定する。また、他のサーバ計算機へ部分ファイルを移動する分散ファイル移動部１１３１、１１３２、１１３３を備えることにより、部分ファイルを他のサーバ計算機に移動することによって、特定のサーバ計算機の記憶装置への負荷の集中を回避できる。

【０１４４】なお、上述した第３の実施の形態では、分散ファイル移動部１１３１の動作アルゴリズム（図１５）のステップ１５０１において、記憶装置の各記憶部の「負荷」が所定の値を越えたことを検知する代わりに、記憶装置の各記憶部の「残容量」が所定の値、例えば、１０［Mbytes］などの値（ただし、この値は、装置やシステムの構成に応じて決定される）を下回ったことを検知するようにしてもよい。これによって、記憶装置の各記憶部の容量の不均衡を回避できる。

【０１４５】また、上述のステップ１５０１において、記憶装置の各記憶部の「負荷」が所定の値を越えたことを検知する代わりに、ネットワーク１０１の「負荷情報」から「使用通信帯域幅」が所定の値、例えば、使用可能通信帯域幅の８０［%］の値（ただし、この値は、装置やシステムの構成に応じて決定される）を越えたリンクを検知するようにすることもできる。また、ステップ１５０２において、アクセス情報テーブル１２０１からネットワーク１０１の負荷を高めている部分ファイルと、ネットワークの負荷を高めている計算機サイトを選択することによって、ネットワークの負荷の集中を回避することができる。例えば、計算機サイトＡ１１０２（送出元サイト）と計算機サイトＢ１１０３（送出先サイト）のリンクの「使用通信帯域幅」（図４）が所定の値を越えた時、サーバ計算機Ａ１１０５に存在し、ネットワーク負荷を高める原因となっている部分ファイルを、サーバ計算機Ｂ１１０３に移動する。これにより、計算機サイトＡ１１０２と計算機サイトＢ１１０３間の「使用通信帯域幅」を減少させることができる。

【０１４６】また、上述の第３の実施の形態では、ステップ１５０２において、移動先のサーバ計算機の部分ファイルを移動可能かを確認し、ステップ１５０３において、部分ファイルの移動を行なっているが、事前にステップ１５０２で部分ファイルの移動可能かを確認をすることなく部分ファイルを移動することによって、ステップ１５０２の確認処理を省略することができる。このとき、部分ファイルの移動先のサーバ計算機側で、部分ファイルの移動を受け入れられない場合には、移動先のサーバ計算機が、さらに部分ファイルを移動するための移動先を探し、この部分ファイルを移動するようにすればよい。

【０１４７】また、上述のステップ１５０３において、部分ファイルを移動元のサーバ計算機から、移動先のサーバ計算機へ移動しているが、その移動処理に加えて、移動先のサーバ計算機内の部分ファイルの中から移動元のサーバ計算機へ移動可能な他の部分ファイルを選択し、当該他の部分ファイルを移動元のサーバ計算機に移動するようにしてもよい。これによって、部分ファイルが１つのサーバ計算機に集中することを防ぐことができ、よりファイルアクセスに対する負荷を軽減することができる。

【０１４８】また、上述のステップ１５０２において、部分ファイルの移動先のサーバ計算機の選択を行なう際に、あらかじめサーバ計算機リストを設定し、リスト中のサーバ計算機の中から、記憶装置の各記憶部の残容量が十分あり、負荷が所定の値より低い記憶装置を持つサーバ計算機を選択するようにするとよい。これによって、部分ファイルの移動先のサーバ計算機の選択に費やされる時間を短縮することができる。

【０１４９】（実施の形態４）図１８は、図１１に示した分散ファイル管理システムの分散ファイル移動部１１３１、１１３２、１１３３の他の動作アルゴリズムを示すフローチャートである。

【０１５０】図１８において、まず、分散ファイル移動部１１３１、１１３２、１１３３は、各部分ファイルへの通信コストを所定の間隔で監視する（ステップ１８０１）。ここで、通信コストとしては、例えば、部分ファイルを参照しているクライアント計算機と、その部分ファイルを保持するサーバ計算機との間の通信時間とすることができる。図１１において、例えば、クライアント計算機１（１０８－１）と部分ファイルＡ２（１２６ｃ）の通信コストは、部分ファイルＡ２（１２６ｃ）を参照しているクライアント計算機１（１０８－１）と、部分ファイルＡ２（１２６ｃ）を保持するサーバ計算機Ｃ１１０７との間の通信時間とする。

【０１５１】ここで、分散ファイル移動部１１３１、１１３２、１１３３は、部分ファイルへの通信コストが所定の値、例えば、１秒など（ただし、この値は、装置やシステムの構成に応じて決定される）を越えたことを検知する（ステップ１８０１）と、通信コストが所定の値を越えた部分ファイルを移動元の部分ファイルとして選択する（ステップ１８０２）。また、この通信コストが所定の値を越えた部分ファイルに対して、複数のクライアント計算機がアクセスしている場合、各々のアクセスに対する通信コストを求め、これらを加算して合計通信コストを求める。

【０１５２】移動先のサーバ計算機を選択する際には、

外部負荷情報テーブル９０１に基づいて、記憶装置の各記憶部の「残容量」が十分にあり、それらに対する「負荷」が所定の値より低い記憶装置を持つサーバ計算機を選択する。そして、選択したサーバ計算機に対して、上述の合計通信コストを送信し、部分ファイルを移動した結果、通信コストがどのように変化するかを順に問い合わせ、最小の通信コストになるサーバ計算機を選択する（ステップ１８０２）。または、分散ファイル移動部１１３１、１１３２、１１３３が、サイト間の接続情報を持ち、その接続情報から通信コストを予想して、最小の通信コストになるサーバ計算機を選択するようにしてもよい（ステップ１８０２）。

【０１５３】図１９は、接続情報テーブルの一例を示す図である。図１９において、接続情報テーブル１９０１は、部分ファイルを送出する「送出元サイト」と、部分ファイルが送出される「送出先サイト」と、送出元サイトから送出先サイトまでの通信コストを示す「通信時間」の項目を有する。この接続情報テーブル１９０１から、部分ファイルを保持するサーバ計算機がどのサイトに属し、部分ファイルを参照するクライアント計算機がどのサイトに属すかにより、サーバ計算機とクライアント計算機の間の「通信時間」を得ることができる。この「通信時間」と上述した合計通信コストに基づいて、部分ファイルの移動後に「通信時間」が最小となるサーバ計算機を、部分ファイルの移動先として選択する（ステップ１８０２）。ここで、通信コストが所定の値を越えた部分ファイルに複数のクライアント計算機がアクセスしている場合には、部分ファイルの移動後の「通信時間」が最小の通信コストになるサーバ計算機を選択するようにするとよい。

【０１５４】例えば、クライアント計算機１（１０８－１）と部分ファイルＡ２（１２６ｃ）の通信コスト（以下、「コストＡ２」ともいう）が所定の値を越えたとき、部分ファイルＡ２（１２６ｃ）をサーバ計算機Ａ１１０５に移動した場合に、通信コストがどのようになるかをサーバ計算機Ａ１１０５に問い合わせるか、または、接続情報テーブル１９０１から通信コストを求める。その結果が、コストＡ２を下回っていれば、サーバ計算機Ａ１１０５を移動先サーバ計算機の候補にする。この処理を他のサーバ計算機に対しても行い、通信コストが最小になるサーバ計算機を探す（ステップ１８０２）。

【０１５５】最後に、この移動元部分ファイルの選択と移動先サーバ計算機の選択（ステップ１８０２）によって得られた情報に基づいて、部分ファイルの移動を行なう（ステップ１８０３）。

【０１５６】以上のように、第４の実施の形態では、分散ファイル移動部１１３１、１１３２、１１３３が、負荷情報テーブル４０１、外部負荷情報テーブル９０１及びアクセス情報テーブル１２０１の各情報と、接続情報テーブル１９０１から得られるサーバ計算機とクライアント計算機の間の接続情報とに基づいて、処理中の部分ファイルが存在するサーバ計算機と処理の要求元のクライアント計算機との間の通信コストを求め、通信コストが所定の値を超えた場合に、通信コストの小さくなる他のサーバ計算機へ部分ファイルを移動するようにしたので、上述した第３の実施の形態で得られる効果に加え、クライアント計算機から部分ファイルへのアクセス時間の平均値を短縮することができる。

【０１５７】なお、上述の第４の実施の形態では、通信コストとして、通信時間を例にあげているが、通信時間の「遅延」や「ゆらぎ（変動幅）」などにすることもできる。

【０１５８】また、上述の第４の実施の形態では、複数のクライアント計算機が同一の部分ファイルにアクセスしている場合、合計通信コストを最小にするように部分ファイルを移動しているが、平均通信コストを最小にするように部分ファイルを移動するようにしてもよい。

【０１５９】（実施の形態５）図２０は、本発明における分散ファイル管理システムの第５の実施の形態の一例を示す構成図である。この図２０においては、図８と同様の構成には同一の符号を付している。図２０に示した分散ファイル管理システムは、パーソナルコンピュータやワークステーションなどのサーバ計算機及びパーソナルコンピュータやワークステーションなどの複数のクライアント計算機から成るクライアント計算機群を備えた複数の計算機サイトＡ２００２、計算機サイトＢ２００３、及び計算機サイトＣ２００４と、計算機サイトＡ２００２、計算機サイトＢ２００３、及び計算機サイトＣ２００４を相互に接続するローカルエリアネットワークやワイドエリアネットワークなどのネットワーク１０１とを備えている。

【０１６０】ここで、計算機サイトＡ２００２は、パーソナルコンピュータやワークステーションなどの複数のサーバ計算機（図２０においては、「サーバ計算機Ａ２００５」のみ示す）と、パーソナルコンピュータやワークステーションなどのクライアント計算機１～ｎ（１０８－１～１０８－ｎ）から成るクライアント計算機群Ａ１０８とを備えている。この計算機サイトＡ２００２は、複数のサーバ計算機（図２０においては、「サーバ計算機Ａ２００５」のみ示す）とクライアント計算機群Ａ１０８とをイーサネットなどの内部ネットワーク１３１で接続しており、例えば、インターネットドメインになっている。

【０１６１】また、計算機サイトＡ２００２と同様に、計算機サイトＢ２００３は、複数のサーバ計算機（図２０においては、「サーバ計算機Ｂ２００６」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｂ１０９とを備え、計算機サイトＣ２００４は、複数のサーバ計算機（図２０においては、「サーバ計算

機Ｃ２００７」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｃ１１０とを備えている。さらに、これらの計算機サイトＢ２００３及び計算機サイトＣ２００４は、計算機サイトＡ２００２と同様に、複数のサーバ計算機（図２０においては、「サーバ計算機Ｂ２００６」及び「サーバ計算機Ｃ２００７」のみ示す）と、クライアント計算機群Ｂ１０９及びクライアント計算機群Ｃ１１０とを、それぞれ内部ネットワーク１３２及び内部ネットワーク１３３で接続しており、例えば、インターネットドメインになっている。

【０１６２】サーバ計算機Ａ２００５は、分散ファイルの部分ファイルを記録するハードディスクなどの記憶装置１１５と、イーサネットなどの内部ネットワーク１３１へ接続するためのネットワークインタフェース１１３と、部分ファイルを記録している記憶装置１１５への書き込みや読み出しを制御する部分ファイル管理部１１１と、記憶装置１１５に対する負荷や記憶装置１１５の残り容量、及びネットワークインタフェース１１３に対する負荷を監視し、これらの負荷や容量に関する情報を保持する状態管理部８１４と、部分ファイル管理部１１１、状態管理部８１４、及びネットワークインタフェース１１３に接続された分散ファイル管理部２０１２とによって構成されている。

【０１６３】この状態管理部８１４は、他のサーバ計算機へ負荷情報を通知し、また、他のサーバ計算機から通知された外部負荷情報を保持する外部状態管理部８１１を備えている。

【０１６４】また、分散ファイル管理部２０１２は、アクセス情報テーブル１２０１（図１２）、負荷情報テーブル４０１（図４）及び外部負荷情報テーブル９０１（図９）から得られる部分ファイル毎の各情報に基づいて、コピーする部分ファイルを決定し、他のサーバ計算機へ当該部分ファイルをコピーする分散ファイルコピー部２００１を備えている。

【０１６５】サーバ計算機Ｂ２００６及びサーバ計算機Ｃ２００７は、サーバ計算機Ａ２００５と同様の構成になっている。すなわち、サーバ計算機Ｂ２００６は、記憶装置１２０と、ネットワークインタフェース１１８と、部分ファイル管理部１１６と、外部状態管理部８１２を備えた状態管理部８１９と、分散ファイルコピー部２０３２を備えた分散ファイル管理部２０１７とによって構成されている。また、サーバ計算機Ｃ２００７は、記憶装置１２５と、ネットワークインタフェース１２３と、部分ファイル管理部１２１と、外部状態管理部８１３を備えた状態管理部８２４と、分散ファイルコピー部２０３３を備えた分散ファイル管理部２０２２とによって構成されている。

【０１６６】ここで、図２０に示した分散ファイル管理システムと図８に示した分散ファイル管理システムとの相違点は、図２０に示した分散ファイル管理部２０１２、２０１７、２０２２が、アクセス情報テーブル１２０１、負荷情報テーブル４０１及び外部負荷情報テーブル９０１から得られる部分ファイル毎の各情報に基づいて、コピーする部分ファイルを決定し、他のサーバ計算機へ当該部分ファイルをコピーする分散ファイルコピー部２０３１、２０３２、２０３３をを備えている点である。

【０１６７】以上のように構成された分散ファイル管理システムの動作について、分散ファイルＡ、分散ファイルＢ及び分散ファイルＣが、サーバ計算機Ａ２００５によって図２０に示すように作成された後に、部分ファイルのコピーを行なう場合を例にとって詳細に説明する。

【０１６８】図２１は、サーバ計算機Ａ２００５の分散ファイルコピー部２０３１の動作アルゴリズムを示すフローチャートである。図２１において、まず、分散ファイルコピー部２０３１は、状態管理部８１４が管理している負荷情報テーブル４０１（図４）を、所定の時間間隔で監視する（ステップ２１０１）。

【０１６９】分散ファイルコピー部２０３１は、記憶装置１１５の任意の記憶部の「負荷」が所定の値、例えば、８０［％］などの値を越えたことを検出する（ステップ２１０１）と、検出された記憶装置１１５の記憶部に含まれている部分ファイルを、部分ファイル管理テーブル１３０１（図１３）を参照して探し出す。そして探し出した部分ファイルのアクセス情報を、アクセス情報テーブル１２０１から獲得する。得られたアクセス情報の「アクセス回数」を比較し、最も大きな「アクセス回数」になっている部分ファイルを、コピー元部分ファイルとして選択する。ここで、例えば、部分ファイルＡ１（１２６ａ）が選択されたとする。次に、外部負荷情報テーブル９０１（図９）に基づいて、記憶装置の記憶部の「残容量」が十分あり、「負荷」が所定の値より低い記憶装置を持つサーバ計算機を選択する。そして、選択したサーバ計算機に対して、部分ファイルがコピーできるかどうかを確認し、コピー可能なサーバ計算機をコピー先サーバ計算機として選択する（ステップ２１０２）。

【０１７０】ここで、例えば、サーバ計算機Ｃ２００７の記憶装置１２５の記憶部（記憶装置識別子：ＤｉｓｋＩＤ２）が選択されたとする。

【０１７１】このコピー元部分ファイルの選択とコピー先サーバ計算機の選択によって得られた情報に基づいて、部分ファイルのコピーを行ない（ステップ２１０３）、再びステップ２１０１の監視処理を続行する。

【０１７２】ここで、上述の例の場合、ステップ２１０２で、コピー元部分ファイルとして部分ファイルＡ１（１２６ａ）、コピー先サーバ計算機としてサーバ計算機Ｃ２００７が選択されたので、コピー元のサーバ計算機Ａ２００５の分散ファイルコピー部２０３１は、部分ファイル管理部１１１を介して、記憶装置１１５から部

分ファイルＡ１（１２６ａ）を読み出し、この部分ファイルＡ１（１２６ａ）をネットワークインタフェース１１３を介してネットワーク１０１へ送出する。また、同時に部分ファイルＡ１（１２６ａ）の「オリジナル所在地」（図１３）に関する情報も送出する。

【０１７３】一方、コピー先のサーバ計算機Ｃ２００７では、分散ファイルコピー部２０３３が、コピー元のサーバ計算機Ａ２００５より送出された部分ファイルＡ１（１２６ａ）を、ネットワークインタフェース１２３を介して受信しする。そして、部分ファイル管理部１２１を介して、記憶装置１２５の所定の記憶部に書き込む。また、部分ファイルＡ１（１２６ａ）の「オリジナル所在地」も受信して、部分ファイル管理テーブル１４０３に登録する。その後、コピー先のサーバ計算機Ｃ２００７は、コピー元のサーバ計算機Ａ２００５と「オリジナル所在地」に示されているサーバ計算機（この例の場合には、「オリジナル所在地」もサーバ計算機Ａ２００５の記憶装置１１５の記憶部である）に対して、部分ファイルＡ１（１２６ａ）のコピーが完了したことを通知する。コピー元のサーバ計算機と「オリジナル所在地」に示されるサーバ計算機（両方ともサーバ計算機Ａ２００５）では部分ファイル管理テーブル１４０１の部分ファイルＡ１（１２６ａ）の情報を書き換える。

【０１７４】図２２は、コピー処理後のサーバ計算機の部分ファイル管理テーブルを示す図である。上述の部分ファイルＡ１（１２６ａ）のコピーの結果、図１４に示した部分ファイル管理テーブル１４０１、１４０２、１４０３は、図２２に示す部分ファイル管理テーブル２２０１、２２０２、２２０３の状態に変化する。すなわち、図２２（Ａ）は、サーバ計算機Ａ２００５の部分ファイル管理テーブル２２０１を示し、（Ｂ）は、サーバ計算機Ｂ２００６の部分ファイル管理テーブル２２０２を示し、（Ｃ）は、サーバ計算機Ｃ２００７の部分ファイル管理テーブル２２０３を示している。また、部分ファイル管理テーブル２２０１には、「部分ファイル識別子」がＡ１、Ａ２、Ａ３、Ｂ１、Ｃ１及びＣ２で示される各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。また、部分ファイル管理テーブル２００２には、「部分ファイル識別子」がＣ１及びＣ２で示される各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。また、部分ファイル管理テーブル２２０３には、「部分ファイル識別子」がＡ１、Ａ２及びＡ３で示される各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。図１４に示した部分ファイル管理テーブルの状態と、図２２に示した部分ファイル管理テーブルの状態の差異は、部分ファイルＡ１（１２６ａ）を、サーバ計算機Ａ２００５からサーバ計算機Ｃ２００７へコピーしたことによる差異である。すなわち、図１４（Ａ）の部分ファイル管理テーブル１４０１において、部分ファイルＡ１（１２６ａ）の「所在地」が、「ｆｉｌｅ：／／ｓｉｔｅＡ／ｓｅｒｖｅｒＡ／ＤｉｓｉＩＤ１／」であるのに対して、図２２（Ａ）の部分ファイル管理テーブル２２０１においては、部分ファイルＡ１（１２６ａ）の「所在地」、が「ｆｉｌｅ：／／ｓｉｔｅＡ／ｓｅｒｖｅｒＡ／ＤｉｓｋＩＤ１／」及び「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／（コピーの所在地）」となっている点が相違する。さらに、図２２（Ｃ）の部分ファイル管理テーブル２１０３においては、部分ファイルＡ１（１２６ａ）の項目が追加されている。

【０１７５】また、図２２の状態から、サーバ計算機Ｂ２００６の分散ファイルコピー部２０３２が、部分ファイルＣ１（１２６ｅ）を、サーバ計算機Ｃ２００７にコピーし、サーバ計算機Ｃ２００７の分散ファイルコピー部２０３３が、部分ファイルＡ１（１２６ａ）を、サーバ計算機Ｂ２００６にコピーすることもできる。

【０１７６】図２３は、図２２の状態から、さらに部分ファイルをコピーした状態の部分ファイル管理テーブルを示す図である。図２３（Ａ）は、サーバ計算機Ａ２００５の部分ファイル管理テーブル２３０１を示し、（Ｂ）は、サーバ計算機Ｂ２００６の部分ファイル管理テーブル２３０２を示し、（Ｃ）は、サーバ計算機Ｃ２００７の部分ファイル管理テーブル２３０３を示している。図２３（Ａ）の部分ファイル管理テーブル２３０１には、「部分ファイル識別子」がＡ１、Ａ２、Ａ３、Ｂ１、Ｃ１及びＣ２の各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。図２３（Ｂ）の部分ファイル管理テーブル２３０２には、「部分ファイル識別子」がＣ１、Ｃ２及びＡ１の各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。また、図２３（Ｃ）の部分ファイル管理テーブル２３０３には、「部分ファイル識別子」がＡ１、Ａ２、Ａ３及びＣ１の各部分ファイルの「所在地」と、「オリジナル所在地」が示されている。図２２の部分ファイル管理テーブルの状態と図２３の部分ファイル管理テーブルの状態の差異は、部分ファイルＡ１（１２６ａ）を、サーバ計算機Ｃ２００７からサーバ計算機Ｂ２００６へコピーしたことと、部分ファイルＣ１（１２６ｅ）をサーバ計算機Ｂ２００６からサーバ計算機Ｃ２００７へコピーしたことによるものである。

【０１７７】すなわち、部分ファイルＡ１（１２６ａ）のコピーに対応して、部分ファイル管理テーブル２３０２において、部分ファイルＡ１（１２６ａ）の項目が追加されている。また、サーバ計算機Ｂ２００６が部分ファイルＡ１（１２６ａ）の「オリジナル所在地」を参照して、サーバ計算機Ａ２００５にコピーを知らせることで、「オリジナル所在地」のサーバ計算機Ａ２００５が、部分ファイル管理テーブル２３０１において、部分ファイルＡ１（１２６ａ）の「所在地」に、「ｆｉｌｅ：／／ｓｉｔｅＢ／ｓｅｒｖｅｒＢ／ＤｉｓｋＩＤ２

／」を追加している。また、部分ファイルＣ１（１２６ｅ）のコピーに対応して、部分ファイル管理テーブル２３０３において、部分ファイルＣ１（１２６ｅ）の項目が追加されている。さらに、部分ファイル管理テーブル２３０２において、部分ファイルＣ１（１２６ｅ）の「所在地」が、「ｆｉｌｅ：／／ｓｉｔｅＢ／ｓｅｒｖｅｒＢ／ＤｉｓｋＩＤ３／」および「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／（コピーの所在地）」となっている。

【０１７８】図２３に示す状態で、クライアント計算機１（１０８－１）が、分散ファイルＣを参照する際に、その参照内容が部分ファイルＣ１（１２６ｅ）に含まれる場合の動作を説明する。

【０１７９】（１）クライアント計算機１（１０８－１）は、分散ファイルＣを作成したサーバ計算機Ａ２００５に対して、分散ファイルＣの参照を要求する。サーバ計算機Ａ２００５では、分散ファイル管理テーブル２３０１を参照して、分散ファイルＣを構成する部分ファイルＣ１、Ｃ２のうちどの部分ファイルを参照しているのかを調べ、部分ファイルＣ１（１２６ｅ）の参照であることを認識する。部分ファイルＣ１（１２６ｅ）の「所在地」は、分散ファイル管理テーブル２３０１では「ｆｉｌｅ：／／ｓｉｔｅＢ／ｓｅｒｖｅｒＢ／ＤｉｓｋＩＤ３／」なので、サーバ計算機Ａ２００５は、サーバ計算機Ｂ２００６に対して、部分ファイルＣ１（１２６ｅ）が存在するかどうかの確認を行なう。

【０１８０】（２）サーバ計算機Ｂ２００６は、分散ファイル管理テーブル２３０２を調べて、部分ファイルＣ１（１２６ｅ）の所在地を調べる。部分ファイルＣ１（１２６ｅ）の「所在地」は、分散ファイル管理テーブル２３０２では「ｆｉｌｅ：／／ｓｉｔｅＢ／ｓｅｒｖｅｒＢ／ＤｉｓｋＩＤ３／」及び「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」なので、サーバ計算機Ｂ２００６は、負荷情報テーブル４０１と外部負荷情報テーブル９０１からサーバ計算機Ｂ２００６とサーバ計算機Ｃ２００７のうち負荷の低いサーバ計算機を選択する。ここで、サーバ計算機Ｃ２００７を選択した場合、サーバ計算機Ｂ２００６は、サーバ計算機Ｃ２００７に対して、部分ファイルＣ１（１２６ｅ）が存在するかどうかの確認を行なう。

【０１８１】（３）サーバ計算機Ｃ２００７は、分散ファイル管理テーブル２３０３を調べて、部分ファイルＣ１（１２６ｅ）の「所在地」を確認する。部分ファイルＣ１（１２６ｅ）の「所在地」は、分散ファイル管理テーブル２３０３では「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」なので、部分ファイルＣ１（１２６ｅ）は、サーバ計算機Ｃ２００７に存在することが解る。

【０１８２】（４）サーバ計算機Ｃ２００７は、サーバ計算機Ｂ２００６に、部分ファイルＣ１（１２６ｅ）が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」に存在することを通知する。

【０１８３】（５）サーバ計算機Ｂ２００６は、サーバ計算機Ａ２００５に、部分ファイルＣ１（１２６ｅ）が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ３／」に存在することを通知する。

【０１８４】（６）サーバ計算機Ａ２００５は、サーバ計算機Ｃ２００７に部分ファイルＣ１（１２６ｅ）の参照を要求する。この要求と同時に、参照を要求したクライアント計算機１（１０８－１）に対して、クライアント計算機１（１０８－１）がサーバ計算機Ｃ２００７に直接に、部分ファイルＣ１（１２６ｅ）の参照要求を行なうように指示する。

【０１８５】以上のように、本実施の形態では、分散ファイル管理部１１２、１１７、１２２が、アクセス情報テーブル１２０１、負荷情報テーブル４０１及び外部負荷情報テーブル９０１から得られる部分ファイル毎の各情報に基づいて、コピーする部分ファイルを決定し、分散ファイルコピー部２０３１、２０３２、２０３３が、部分ファイルを他のサーバ計算機にコピーすることによって、特定のサーバ計算機の記憶装置への負荷の集中を回避することができる。

【０１８６】なお、上述した第５の実施の形態では、分散ファイルコピー部２０３１の動作アルゴリズム（図２１）のステップ２１０１において、記憶装置の各記憶部の「負荷」が所定の値を越えたことを検知するのに代えて、ネットワーク負荷情報テーブル４０３の「使用通信帯域幅」が所定の値を越えたリンクを検知することにしてもよい。また、ステップ２１０２において、アクセス情報テーブル１２０１からネットワーク負荷を高めている部分ファイルと、ネットワーク負荷を高めている計算機サイトを選択するようにしてもよい。これによって、ネットワーク負荷の集中を回避できる。例えば、計算機サイトＡ２００２（送出元サイト）と計算機サイトＢ２００３（送出先サイト）のリンクの「使用通信帯域幅」が所定の値を越えた時、サーバ計算機Ａ２００５に存在し、ネットワーク負荷を高める原因となっている部分ファイルをサーバ計算機Ｂ２００３にコピーする。これにより、計算機サイトＡ２００２と計算機サイトＢ２００３間の「使用通信帯域幅」が減少する。

【０１８７】図２４は、サーバ計算機Ａ２００５の分散ファイルコピー部２０３１の他の動作アルゴリズムを示すフローチャートである。図２４において、まず、分散ファイルコピー部２０３１は、部分ファイルへの通信コストを所定の間隔で監視する（ステップ２４０１）。

【０１８８】ここで、通信コストには、例えば、部分ファイルを参照しているクライアント計算機と、その部分ファイルを保持するサーバ計算機との間の通信時間とすることができる。例えば、クライアント計算機１（１０８－１）と部分ファイルＡ２（１２６ｃ）の通信コスト

の場合は、部分ファイルＡ２（１２６ｃ）を参照しているクライアント計算機１（１０８－１）と、部分ファイルＡ２（１２６ｃ）を保持するサーバ計算機Ｃ２００７の間の通信時間とすればよい。

【０１８９】次に、分散ファイルコピー部２０３１は、部分ファイルへの通信コストが所定の値を越えたことを検知した場合（ステップ２４０１）、この通信コストが所定の値を越えた部分ファイルを、コピー元部分ファイルとして選択する。通信コストが所定の値を越えた部分ファイルに、複数のクライアント計算機がアクセスしている場合には、各アクセス毎に各々の通信コストを求め、これらを加算して合計通信コストを求める。一方、コピー先のサーバ計算機を選択する際には、外部負荷情報テーブル９０１に基づいて、記憶装置の記憶部の「残容量」が十分あり、「負荷」が低い記憶装置を持つサーバ計算機を選択し、選択されたサーバ計算機に対して合計通信コストを送信し、部分ファイルをコピーした結果、通信コストがどのように変化するかを順に問い合わせる。そして、最小の通信コストになるサーバ計算機をコピー先サーバ計算機として選択する。または、分散ファイルコピー部２０３１、２０３２、２０３３が、サイト間（サーバ計算機とクライアント計算機との間）の接続情報として、例えば、図１９に示す接続情報テーブル１９０１を持ち、その情報から通信コスト（通信時間）を予想して、最小の通信コストになるサーバ計算機を選択するようにしてもよい。図１９の接続情報テーブル１９０１の情報と、サーバ計算機がどのサイトに属しクライアント計算機がどのサイトに属すかの情報とにより、サーバ計算機とクライアント計算機の間の「通信時間」を得ることができる。そして、その「通信時間」が最小となるサーバ計算機を、部分ファイルのコピー先として選択するとよい。通信コストが所定の値を越えた部分ファイルに、複数のクライアント計算機がアクセスしている場合には、最小の合計通信コストになるサーバ計算機をコピー先サーバ計算機として選択する（ステップ２４０２）。

【０１９０】例えば、クライアント計算機１（１０８－１）と部分ファイルＡ２の通信コスト（コストＡ２）が所定の値を越えた時には、部分ファイルＡ２をサーバ計算機Ａ１０５にコピーした場合に、通信コストがどのようになるかをサーバ計算機Ａ１０５に問い合わせる。あるいは、接続情報１８０１からコストを求める。その結果が、コストＡ２を下回っていれば、サーバ計算機Ａ１０５をコピー先の候補にする。この処理を繰り返して、通信コストが最小になるサーバ計算機を探す。このコピー元部分ファイルの選択とコピー先サーバ計算機の選択によって得られた情報に基づいて、部分ファイルのコピーを行ない（ステップ２４０３）、再びステップ２４０１の監視処理を続行する。

【０１９１】このように、図２１のステップ２１０１とステップ２１０２を図２４のステップ２４０１とステップ２４０２に変更することで、クライアント計算機から部分ファイルへのアクセス時間の平均値を短縮することができる。また、上記では通信コストとして、「通信時間」を例にあげているが、通信時間の「遅延」や「ゆらぎ（変動幅）」などでもよい。

【０１９２】また、図２１のステップ２１０２及び図２４のステップ２４０２において、コピー先のサーバ計算機に部分ファイルのコピーが可能かを確認し、ステップ２１０３及びステップ２４０３において、部分ファイルのコピーを行なっているが、ステップ２１０２及びステップ２４０２でサーバ計算機に部分ファイルのコピーが可能かを確認することなく、部分ファイルをコピーすることによって、ステップ２１０２及びステップ２４０２の確認処理を省略することができる。このとき、ステップ２１０３及びステップ２４０３において、コピー先のサーバ計算機側で、部分ファイルのコピーを受け入れられない場合には、コピー先のサーバ計算機は、さらに部分ファイルをコピーするためのコピー先を探して部分ファイルをコピーするか、あるいは、コピー用に送信されてきた部分ファイルを破棄して、コピー元のサーバ計算機にコピー用の部分ファイルを破棄したことを通知するようにしてもよい。

【０１９３】また、上述の第５の実施の形態では、図２１のステップ２１０３及び図２４のステップ２４０３において、部分ファイルをコピー元のサーバ計算機から、コピー先のサーバ計算機へ単にコピーしているが、そのコピー処理に加えて、コピー先のサーバ計算機内の部分ファイルの中からコピー元のサーバ計算機へ移動させても良い部分ファイルを選択し、その部分ファイルを移動元のサーバ計算機に移動するようにしてもよい。これによって、１つのサーバ計算機に部分ファイルが集中せず、特定のサーバ計算機の記憶装置への負荷の集中を回避することができる。

【０１９４】また、図２１のステップ２１０２において、コピー先のサーバ計算機の選択を行なう際に、あらかじめサーバ計算機リストを設定しておき、このサーバ計算機リスト中のサーバ計算機の中から、記憶装置の記憶部の残容量が十分あり、負荷が低い記憶装置を持つサーバ計算機を選択するようにすることもできる。これによって、サーバ選択のための時間を短縮することができるようになる。

【０１９５】（実施の形態６）図２５は、図２０に示した分散ファイル管理システムにおけるサーバ計算機Ａ２００５の分散ファイルコピー部２０３１の他の動作アルゴリズムを示すフローチャートである。本実施の形態においては、上述した第５の実施の形態と同様の構成となっており、分散ファイルコピー部の動作を除いては同様の動作を行う。以下、分散ファイルコピー部２０３１の動作アルゴリズムについて説明する。

【０１９６】図２５において、まず、分散ファイルコピー部２０３１は、状態管理部８１４が管理している負荷情報テーブル４０１（図４）を、所定の時間間隔で監視する（ステップ２５０１）。

【０１９７】分散ファイルコピー部２０３１は、記憶装置１１５の任意の記憶部の「負荷」が所定の値、例えば、８０［％］などの値を越えたことを検出する（ステップ２５０１）と、検出された記憶装置に含まれている部分ファイルを、部分ファイル管理テーブル１３０１（図１３）を参照して探し出す。そして、探し出した部分ファイルの「単位時間当たりのアクセス情報」を、アクセス情報テーブル１２０１から獲得する。獲得した「単位時間当たりのアクセス情報」の「アクセス回数」を比較し、最も大きな「アクセス回数」になっている部分ファイルを、コピー元部分ファイルとして選択する（ステップ２５０２）。ここでは、例えば、部分ファイルＡ１が選択されたとする。

【０１９８】次に、外部負荷情報テーブル９０１（図９）に基づいて、記憶装置の記憶部の「残容量」が十分あり、「負荷」が所定の値より低い記憶装置を持つ複数のサーバ計算機を選択する。そして、選択されたサーバ計算機に対して部分ファイルをコピーできるかどうかを確認し、コピー可能なサーバ計算機を、コピー先サーバ計算機として決定する（ステップ２５０２）。

【０１９９】ここで、例えば、サーバ計算機Ｂ１０６の記憶装置識別子ＤｉｓｋＩＤ１で示される記憶部を有する記憶装置１２０と、サーバ計算機Ｃ１０７の記憶装置識別子ＤｉｓｋＩＤ２で示される記憶部を有する記憶装置１２５が選択されたとする。

【０２００】次に、上述のステップ２５０２で得られたコピー元部分ファイルとコピー先サーバ計算機の情報に基づいて、部分ファイルのコピーを行ない（ステップ２５０３）、ステップ２５０１の監視処理を続行する。

【０２０１】上述の例において、ステップ２５０２で、コピー元部分ファイルとして部分ファイルＡ１（１２６ａ）、コピー先計算機としてサーバ計算機Ｂ２００６とサーバ計算機Ｃ２００７が選択されたので、コピー元のサーバ計算機Ａ２００５の分散ファイルコピー部２０３１は部分ファイル管理部１１１を介して、記憶装置１１５から部分ファイルＡ１（１２６ａ）を読み出し、ネットワークインタフェース１１３を介してネットワーク１０１へマルチキャストで送信する。また、同時に部分ファイルＡ１（１２６ａ）の「オリジナル所在地」（図１３）に関する情報も送信する。

【０２０２】一方、コピー先のサーバ計算機Ｂ２００６と計算機Ｃ２００７では、分散ファイルコピー部２０３２、２０３３がコピー元のサーバ計算機Ａ２００５より送信された部分ファイルＡ１（１２６ａ）を、ネットワークインタフェース１１８、１２３を介して受信し、部分ファイル管理部１１６、１２１を介して、記憶装置１２０、１２５に書き込む。また、部分ファイルＡ１（１２６ａ）の「オリジナル所在地」（図１３）も受信して、それぞれの部分ファイル管理テーブル１４０２、１４０３（図１４）に登録する。その後、コピー先のサーバ計算機Ｂ２００６とサーバ計算機Ｃ２００７は、それぞれコピー元のサーバ計算機と「オリジナル所在地」（図１３）に示されているサーバ計算機、すなわち、この場合には、両方ともサーバ計算機Ａ２００５であるため、両者ともこのサーバ計算機Ａ２００５に対して、部分ファイルＡ１（１２６ａ）のコピーが完了したことを通知する。コピー元のサーバ計算機と「オリジナル所在地」（図１３）に示されるサーバ計算機、すなわち、サーバ計算機Ａ２００５では部分ファイル管理テーブル１４０１（図１４）の部分ファイルＡ１（１２６ａ）の情報を書き換える。

【０２０３】以上のように、本実施の形態では、分散ファイルコピー部２０３１、２０３２、２０３３が部分ファイルを他のサーバ計算機へコピーする際に、複数のコピー先のサーバ計算機の候補を選択し、選択された複数のサーバ計算機へマルチキャスト通信により同時に部分ファイルをコピーするため、部分ファイルのコピーの際の通信量を削減できる。

【０２０４】（実施の形態７）図２６は、本発明における分散ファイル管理システムの他の実施の形態の一例を示す構成図である。ここで、図２６では、図１と同一の構成のものには同一の符号を付している。図２６において、この分散ファイル管理システムは、パーソナルコンピュータやワークステーションなどのサーバ計算機及びパーソナルコンピュータやワークステーションなどの複数のクライアント計算機から成るクライアント計算機群を備えた複数の計算機サイトＡ２６０２、計算機サイトＢ２６０３、及び計算機サイトＣ２６０４と、計算機サイトＡ２６０２、計算機サイトＢ２６０３、及び計算機サイトＣ２６０４を相互に接続するローカルエリアネットワークやワイドエリアネットワークなどのネットワーク１０１とを備えている。

【０２０５】ここで、計算機サイトＡ２６０２は、パーソナルコンピュータやワークステーションなどの複数のサーバ計算機（図２６においては、「サーバ計算機Ａ２６０５」のみ示す）と、パーソナルコンピュータやワークステーションなどのクライアント計算機１～ｎ（１０８－１～１０８－ｎ）から成るクライアント計算機群Ａ１０８とを備えている。この計算機サイトＡ２６０２は、複数のサーバ計算機（図２６においては、「サーバ計算機Ａ２６０５」のみ示す）とクライアント計算機群Ａ１０８とをイーサネットなどの内部ネットワーク１３１で接続しており、例えば、インターネットドメインになっている。

【０２０６】また、計算機サイトＡ２６０２と同様に、計算機サイトＢ２６０３は、複数のサーバ計算機（図２

６においては、「サーバ計算機Ｂ２６０６」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｂ２６０９とを備え、計算機サイトＣ２６０４は、複数のサーバ計算機（図２６においては、「サーバ計算機Ｃ２６０７」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｃ２６１０とを備えている。さらに、これらの計算機サイトＢ２６０３及び計算機サイトＣ２６０４は、計算機サイトＡ２６０２と同様に、複数のサーバ計算機（図２６においては、「サーバ計算機Ｂ２６０６」及び「サーバ計算機Ｃ２６０７」のみ示す）と、クライアント計算機群Ｂ１０９及びクライアント計算機群Ｃ１１０とを、それぞれ内部ネットワーク１３２及び内部ネットワーク１３３で接続しており、例えば、インターネットドメインになっている。

【０２０７】サーバ計算機Ａ２６０５は、分散ファイルの部分ファイルを記録するハードディスクなどの記憶装置１１５と、イーサネットなどの内部ネットワーク１３１へ接続するためのネットワークインタフェース１１３と、部分ファイルを記録している記憶装置１１５への書き込みや読み出しを制御する部分ファイル管理部１１１と、記憶装置１１５に対する負荷や記憶装置１１５の残り容量、及びネットワークインタフェース１１３に対する負荷を監視し、これらの負荷や容量に関する情報を保持する状態管理部１１４と、部分ファイル管理部１１１、状態管理部１１４、及びネットワークインタフェース１１３に接続された分散ファイル管理部２６１２とによって構成されている。

【０２０８】この分散ファイル管理部２６１２は、部分ファイルの書き込みや読み出しを部分ファイル管理部１１１に指示する。また、分散ファイル管理部２６１２は、分散ファイルを作成する場合には、状態管理部１１４から得られる情報に基づいて、分散ファイルを複数の部分ファイルに分割し、各部分ファイルを配置（記録）するサーバ計算機を決定する。また、以前に作成された分散ファイルを参照または更新する場合には、該当する分散ファイルの部分ファイルが存在する（記録されている）サーバ計算機を検出する。ここで、分散ファイル管理部２６１２は、クライアント計算機からの情報または分散ファイルに記録されるデータの種類に応じて、分散ファイルを部分ファイルへ分割する際の部分ファイルのサイズを決定する部分ファイルサイズ決定部２６３１を備えている。

【０２０９】サーバ計算機Ｂ２６０６及びサーバ計算機Ｃ２６０７は、サーバ計算機Ａ２６０５と同様の構成になっている。すなわち、サーバ計算機Ｂ２６０６は、記憶装置１２０と、ネットワークインタフェース１１８と、部分ファイル管理部１１６と、状態管理部１１９と、分散ファイル管理部２６１７とによって構成されている。また、サーバ計算機Ｃ２６０７は、記憶装置１２５と、ネットワークインタフェース１２３と、部分ファイル管理部１２１と、状態管理部１２４と、分散ファイル管理部２６２２とによって構成されている。また、分散ファイル管理部２６１７、２６２２は、それぞれ部分ファイルサイズ決定部２６３２、２６３３を備えている。

【０２１０】ここで、図２６に示した分散ファイル管理システムと図に示した分散ファイル管理システムとの相違点は、図２６の分散ファイル管理システムにおいて、分散ファイル管理部２６１２、２６１７、２６２２に、クライアント計算機からの情報または分散ファイルに記録されるデータの種類に応じて、分散ファイルを部分ファイルへ分割する際の部分ファイルのサイズを決定する部分ファイルサイズ決定部２６３１、２６３２、２６３３を備えている点である。

【０２１１】以上のように構成された分散ファイル管理システムの部分ファイルサイズ決定部２６３１、２６３２、２６３３の動作について以下に説明する。

【０２１２】図３において、例えば、クライアント計算機１（１０８－１）がサーバ計算機Ａ２６０５に対して、分散ファイルＡの作成要求を行なった時、ステップ３０２の処理で、分散ファイル管理部２６１２の部分ファイルサイズ決定部２６３１が、部分フィルを割り当てる時のサイズを決定する。部分ファイルサイズ決定部２６３１が、部分ファイルのサイズを決定するに際には、分散ファイルに記録するデータの種類（例えば、Ｍ－ＪＰＥＧ、ＭＰＥＧ１、ＭＰＥＧ２など）やクライアント計算機１（１０８－１）からの指示によって決めるようにするとよい。

【０２１３】以上のように、本実施の形態の分散ファイル管理システムにおいて、分散ファイル管理部２６１２、２６１７、２６２２の部分ファイルサイズ決定部２６３１、２６３２、２６３３が、クライアント計算機からの情報や分散ファイルに記録されるデータの種類などによって、分散ファイルを部分ファイルへ分割する際の部分ファイルのサイズを決定することにより、分散ファイルを構成する部分ファイルのサイズを適宜変更することができる。これにより、論理的や内容的に関連のあるデータ、例えば、画像１フレーム分のデータなどを複数の部分ファイルに分割してしまうことを防止することができる。

【０２１４】（実施の形態８）次に、上述した複数のサーバ計算機が有する分散ファイル管理部及び状態管理部を１つのサーバ計算機にまとめ、該サーバ計算機で集中して管理する場合について説明する。

【０２１５】図２７は、本発明における第１の実施の形態に示した分散ファイル管理システムにおいて、分散ファイル管理部及び状態管理部を１つのサーバ計算機にまとめた場合の分散管理システムの構成の一例を示している。図２７において、図１と同一の構成のものには同一

の符号を付している。

【０２１６】図２７に示した分散ファイル管理システムは、パーソナルコンピュータやワークステーションなどのサーバ計算機及びパーソナルコンピュータやワークステーションなどの複数のクライアント計算機から成るクライアント計算機群を備えた複数の計算機サイトＡ２７０２、計算機サイトＢ２７０３、及び計算機サイトＣ２７０４と、他のサーバ計算機上に配置されている分散ファイルを集中的に管理する管理サーバ計算機Ｘ２７１１を備えた計算機サイトＸ２７１０と、計算機サイトＡ２７０２、計算機サイトＢ２７０３、計算機サイトＣ２７０４、及び計算機サイトＸ２７１０を相互に接続するローカルエリアネットワークやワイドエリアネットワークなどのネットワーク１０１とを備えている。

【０２１７】ここで、計算機サイトＡ２７０２は、パーソナルコンピュータやワークステーションなどの複数のサーバ計算機（図２７においては、「サーバ計算機Ａ２７０５」のみ示す）と、パーソナルコンピュータやワークステーションなどのクライアント計算機１～ｎ（１０８－１～１０８－ｎ）から成るクライアント計算機群Ａ１０８とを備えている。この計算機サイトＡ２７０２は、複数のサーバ計算機（図２７においては、「サーバ計算機Ａ２７０５」のみ示す）とクライアント計算機群Ａ１０８とをイーサネットなどの内部ネットワーク１３１で接続しており、例えば、インターネットドメインになっている。

【０２１８】また、計算機サイトＡ２７０２と同様に、計算機サイトＢ２７０３は、複数のサーバ計算機（図２７においては、「サーバ計算機Ｂ２７０６」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｂ１０９とを備え、計算機サイトＣ２７０４は、複数のサーバ計算機（図２７においては、「サーバ計算機Ｃ２７０７」のみ示す）と、複数のクライアント計算機から成るクライアント計算機群Ｃ１１０とを備えている。さらに、これらの計算機サイトＢ２７０３及び計算機サイトＣ２７０４は、計算機サイトＡ２７０２と同様に、複数のサーバ計算機（図２７においては、「サーバ計算機Ｂ２７０６」及び「サーバ計算機Ｃ２７０７」のみ示す）と、クライアント計算機群Ｂ１０９及びクライアント計算機群Ｃ１１０とを、それぞれ内部ネットワーク１３２及び内部ネットワーク１３３で接続しており、例えば、インターネットドメインになっている。

【０２１９】サーバ計算機Ａ２７０５は、分散ファイルの部分ファイルを記録するハードディスクなどの記憶装置１１５と、イーサネットなどの内部ネットワーク１３１へ接続するためのネットワークインタフェース１１３と、部分ファイルを記録している記憶装置１１５への書き込みや読み出しを制御する部分ファイル管理部１１１とによって構成されている。

【０２２０】サーバ計算機Ｂ２７０６及びサーバ計算機Ｃ２７０７は、サーバ計算機Ａ２７０５と同様の構成になっている。すなわち、サーバ計算機Ｂ２７０６は、記憶装置１２０と、ネットワークインタフェース１１８と、部分ファイル管理部１１６とによって構成されている。また、サーバ計算機Ｃ２７０７は、記憶装置１２５と、ネットワークインタフェース１２３と、部分ファイル管理部１２１とによって構成されている。

【０２２１】管理サーバ計算機Ｘ２７１１は、イーサネットなどの内部ネットワーク１３４に接続するためのネットワークインタフェース２７１３と、各サーバ計算機の記憶装置の負荷や残り容量、ネットワークインタフェースの負荷を監視し、負荷に関する情報を保持する状態管理部２７１４と、部分ファイルの書き込みや読み出しを各サーバ計算機の部分ファイル管理部１１１、１１６、１２１に指示したり、分散ファイルを作成する際に、状態管理部２７１４からの情報に基づいて分散ファイルを複数の部分ファイルに分割し、各部分ファイルを配置するサーバ計算機を決定することによって部分ファイルの作成を行ない、また、分散ファイルを参照または更新する場合には、参照または更新される当該分散ファイルの部分ファイルが在左するサーバ計算機を検出して部分ファイルの参照または更新を行なう分散ファイル管理部２７１２によって構成されている。

【０２２２】図２７においては、各分散ファイルＡ、Ｂ、Ｃが作成された後の状態を示している。すなわち、サーバ計算機Ａ２７０５の記憶装置１１５には、分散ファイルＡの部分ファイルＡ１（１２６ａ）と分散ファイルＢの部分ファイルＢ１（１２６ｂ）とが記録されている。また、サーバ計算機Ｂ２７０６の記憶装置１２０には、分散ファイルＣの部分ファイルＣ１（１２６ｅ）と分散ファイルＣの部分ファイルＣ２（１２６ｆ）とが記録されている。また、サーバ計算機Ｃ２７０７の記憶装置１２５には、分散ファイルＡの部分ファイルＡ２（１２６ｃ）と分散ファイルＡの部分ファイルＡ３（１２６ｄ）とが記録されている。

【０２２３】次に、以上のように構成された分散ファイル管理システムの動作について説明する。以下においては、クライアント計算機群Ａ１０８のクライアント計算機１（１０８－１）からサーバ計算機Ａ２７０５に対して分散ファイルＡの作成要求が発行され、図２７に示したような部分ファイルＡ１～Ａ３が作成される場合の分散処理を例にして説明する。ここで、図２７に示した記憶装置１１５、１２０、１２５は、それぞれ複数の記憶部または記憶領域（以下、単に「記憶部」ともいう）を有するものとする。これらの複数の記憶部は、物理的に１つの記録媒体であってもよく、また、複数の記録媒体であってもよい。

【０２２４】図２７において、まず、クライアント計算機１（１０８－１）から計算機サイトＸ２７１０の管理サーバ計算機Ｘ２７１１に分散ファイルＡの作成要求が

発行される。この分散ファイルＡの作成要求は、計算機サイトＡ２７０２の内部ネットワーク１３１、ネットワーク１０１、計算機サイトＸ２７１０の内部ネットワーク１３４、及び管理サーバ計算機Ａ２７１１のネットワークインタフェース２７１３を介して、分散ファイル管理部２７１２によって受け取られる。

【０２２５】図２８は、分散ファイルの作成要求を受け取った場合の分散ファイル管理部２７１２の動作アルゴリズムを示すフローチャートである。以下、図２７及び図２８を用いて、分散ファイル管理部２７１２の詳細な動作を説明する。

【０２２６】図２８において、まず、分散ファイル管理部２７１２は、状態管理部２７１４の管理している負荷情報を参照する（ステップ２８０１）。

【０２２７】状態管理部２７１４では、例えば、図４のような負荷情報テーブル４０１を管理している。図４では、サーバ計算機Ａ２７０５に関する情報について示しているが、状態管理部２７１４では、図４に示したような負荷情報テーブル４０１を各サーバ計算機毎に準備し管理してる。図４において、負荷情報テーブル４０１は、記憶装置負荷情報テーブル４０２とネットワーク負荷情報テーブル４０３からなる。記憶装置負荷情報テーブル４０２は、サーバ計算機に接続されている記憶装置の各記憶部を識別するための「記憶装置識別子」と、記憶装置の各記憶部の「負荷」と、記憶装置の各記憶部の「残容量」の情報で構成されている。記憶装置の各記憶部の「負荷」は、記憶部の最大転送レートのうち、何［％］を使用しているかで表示している。ネットワーク負荷情報テーブル４０３は、それぞれのサーバ計算機のネットワークインタフェースを介して送信するデータがどのサイトに向けてのものであり、どの程度の帯域幅を使用して送信されているか、また、受信しているデータがどのサイトから送られて来たものであり、どの程度の帯域幅を使用して受信しているかを表している。また、「送出元サイト」がデータの送出元の計算機サイトを示し、「送出先サイト」がデータの送出先の計算機サイトを示し、「使用帯域幅」が送出元の計算機サイトと送出先の計算機サイトの間で使用されている通信帯域幅を示している。

【０２２８】例えば、分散ファイルＡを作成する場合、分散ファイル管理部２７１２は、サーバ計算機Ａ２７０５の記憶装置１１５の記憶装置識別子がＤｉｓｋＩＤ１で示される記憶部について、その「負荷」が２０［％］で、その「残容量」が１０［Mbytes］であるという情報を得ることができる。

【０２２９】次に、分散ファイル管理部２７１２は、状態管理部２７１４から得られた「記憶装置負荷情報」に基づいて、各サーバ計算機に接続されている記憶装置の各記憶部の中から、「残容量」の値が所定の値より大きく、「負荷」の値が所定の値、例えば、８０［％］（この値は、システムや他の装置の構成に応じて決定される）より低い記憶部を有する記憶装置を選択する。そして、この記憶装置の記憶部に部分ファイルを順番に割り当てる。このとき、部分ファイルのサイズを全てのサーバ計算機で同一の固定長とするとよい。この部分ファイルの割り当て処理において、全ての部分ファイルを割り当てることができたかどうかを検知する（ステップ２８０２）。

【０２３０】ここで、分散ファイルＡを作成する場合、図２７においては、部分ファイルＡ１（１２６ａ）をサーバ計算機Ａ２７０５に割り当て、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）をサーバ計算機Ｃ２７０４に割り当てることになる。

【０２３１】全ての部分ファイルを割り当てることができた場合には、分散ファイル管理部２７１２は、分散ファイルを管理するための情報を、例えば、上述したような図５の分散ファイル管理テーブル５０１と図６の部分ファイル管理テーブル６０１に登録する（ステップ２８０３）。

【０２３２】上述の例においては、図５に示すように、分散ファイルＡは、部分ファイルＡ１（１２６ａ）、部分ファイルＡ２（１２６ｃ）、及び部分ファイルＡ３（１２６ｄ）から構成されている。

【０２３３】また、図６において、部分ファイルＡ１（１２６ａ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＡ／ｓｅｒｖｅｒＡ／ＤｉｓｋＩＤ１／（計算機サイトＡ１０２のサーバ計算機Ａ１０５の記憶装置識別子ＤｉｓｋＩＤ１）」であり、部分ファイルＡ２（１２６ｃ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／（計算機サイトＣ１０４のサーバ計算機Ｃ１０７の記憶装置識別子ＤｉｓｋＩＤ２）」であり、部分ファイルＡ３（１２６ｄ）の所在地が、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／（計算機サイトＣ１０４のサーバＣ１０７の記憶装置ＤｉｓｋＩＤ２）」であることを表している。

【０２３４】次に、分散ファイル管理部２７１２は、ステップ２８０３で登録した部分ファイルを該当する各サーバ計算機上に作成するため、作成を行なうサーバ計算機の部分ファイル管理部に部分ファイルの作成要求を行う。この作成要求と同時に、分散ファイルの作成要求を行なったクライアント計算機１（１０８－１）に対して、クライアント計算機１（１０８－１）から直接部分ファイルの作成を行なうサーバ計算機にデータを送信するように指示する。部分ファイルの作成を要求されたサーバ計算機では、部分ファイル管理部によって記憶装置の各記憶部にクライアント計算機１（１０８－１）からのデータの書き込みを行なう。分散ファイル管理部２７１２は、全ての部分ファイルの作成が終るまでこの処理を繰り返す（ステップ２８０４）。

【０２３５】ここで、分散ファイルＡの部分ファイルＡ

１（１２６ａ）の作成の場合、サーバ計算機Ａ２７０５の部分ファイル管理部１１１が、クライアント計算機１（１０８－１）からのデータを記憶装置１１５の所定の記憶部に書き込む。また、サーバ計算機Ｃ２７０７の部分ファイル管理部１２１は、クライアント計算機１（１０８－１）からのデータを記憶装置１２５の所定の記憶部に書き込んで、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）を作成する。

【０２３６】一方、ステップ２８０２で、全ての部分ファイルを割り当てることができなかった場合には、分散ファイル管理部２７１２は、分散ファイルの作成要求を行なったクライアント計算機１（１０８－１）に対して、分散ファイルの作成処理が失敗したことを通知する（ステップ２８０５）。

【０２３７】次に、クライアント計算機から管理サーバ計算機に対して分散ファイルの参照または更新の要求（以下、「参照／更新要求」ともいう）が発行された場合について説明する。また、以下の説明において、クライアント計算機群Ａ１０８内のクライアント計算機１（１０８－１）から管理サーバ計算機Ｘ２７１１に対して分散ファイルＡの参照／更新要求が発行された場合を例にして述べる。

【０２３８】まず、クライアント計算機１（１０８－１）によって発行された分散ファイルＡに対する参照／更新要求は、管理サーバ計算機Ｘ２７１１において、ネットワークインタフェース２７１３を介して、分散ファイル管理部２７１２によって受信される。

【０２３９】図２９は、分散ファイル管理部が分散ファイルの参照／更新要求を受け取った場合の動作アルゴリズムを示すフローチャートである。以下、図２９を用いて、分散ファイル管理部２７１２の動作を説明する。

【０２４０】まず、分散ファイル管理部２７１２は、クライアント計算機１（１０８－１）からの分散ファイルＡの参照／更新要求に応じて、分散ファイル管理テーブル５０１（図５）と部分ファイル管理テーブル６０１（図６）から、更新または参照する部分ファイルと、その部分ファイルの所在地を求める（ステップ２９０１）。

【０２４１】ここで、分散ファイルＡの参照／更新要求の場合、分散ファイル管理テーブル５０１（図５）から、分散ファイルＡは、部分ファイルＡ１（１２６ａ）、部分ファイルＡ２（１２６ｃ）及び部分ファイルＡ３（１２６ｄ）により構成されていることがわかる。また、部分ファイル管理テーブル６０１（図６）によって、部分ファイルＡ１（１２６ａ）は、「ｆｉｌｅ：／／ｓｉｔｅＡ／ｓｅｒｖｅｒＡ／ＤｉｓｋＩＤ１／」で示される記憶装置１１５の記憶部に存在し、部分ファイルＡ２（１２６ｃ）は、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／」で示される記憶装置１２５の記憶部に存在し、部分ファイルＡ３（１２６ｄ）は、「ｆｉｌｅ：／／ｓｉｔｅＣ／ｓｅｒｖｅｒＣ／ＤｉｓｋＩＤ２／」で示される記憶装置１２５の記憶部に存在することが解る。

【０２４２】分散ファイル管理部２７１２は、クライアント計算機１（１０８－１）からの参照／更新要求に対応する部分ファイルを保持するサーバ計算機に対して、当該部分ファイルの参照／更新要求を行う。この要求と同時に、分散ファイル管理部２７１２は、クライアント計算機１（１０８－１）に対して、クライアント計算機１（１０８－１）が参照または更新を行う部分ファイルの存在するサーバ計算機に直接、参照／更新要求を行うように指示する。各サーバ計算機の部分ファイル管理部は、分散ファイル管理部２７１２からの要求に応じ、クライアント計算機１（１０８－１）からの参照／更新要求に基づいて、記憶装置に存在する部分ファイルの読み出し（参照）、または記憶装置への部分ファイルの書き込み（更新）を行なう（ステップ２９０２）。

【０２４３】ここで、分散ファイルＡの場合、部分ファイルＡ１（１２６ａ）は、計算機サイトＡ２７０２のサーバ計算機Ａ２７０５に、部分ファイルＡ２（１２６ｃ）は、計算機サイトＣ２７０４のサーバ計算機Ｃ２７０７に、部分ファイルＡ３（１２６ｄ）は、計算機サイトＣ２７０４のサーバ計算機Ｃ２７０７に存在している。したがって、部分ファイルＡ１（１２６ａ）に対する参照／更新要求の処理は、分散ファイル管理部２７１２からの要求に応じて、クライアント計算機１（１０８－１）と部分ファイル管理部１１１との間で直接行なわれる。一方、部分ファイルＡ２（１２６ｃ）と部分ファイルＡ３（１２６ｄ）に対する参照／更新要求の処理は、分散ファイル管理部２７１２からの要求に応じて、クライアント計算機１（１０８－１）と、サーバ計算機Ｃ２７０７との間で直接行なわれる。

【０２４４】以上のように、本実施の形態の分散ファイル管理システムによれば、第１の実施の形態に示した効果に加え、分散ファイルの管理とシステムの状態の管理を集中して行うため、重複した管理部を複数持つ必要がなく、システム構成を簡略することができ、またコストの軽減を図ることができる。

【０２４５】尚、上述した図２７の分散ファイル管理システムにおいては、第１の実施の形態で示した分散ファイル管理システムの管理部を、１つの管理サーバ計算機に集中した構成として説明したが、第２～第７で示した実施の形態の分散ファイル管理システムにも適用することができる。

【０２４６】また、上述した図２７の分散ファイル管理システムにおいて、分散ファイル管理システムの管理部を、１つの管理サーバ計算機に集中した構成として説明したが、所定のグループのサーバ計算機毎や所定のグループの計算機サイト毎に管理サーバ計算機を設けるようにしてもよい。このようにすると、大規模なシステムに

おける管理サーバ計算機への負荷の集中を防止することができる。

【０２４７】また、上述した実施の形態において、部分ファイルが複数のサーバ計第機上に存在する場合の部分ファイルの選択においては、単にサーバ計算機の「負荷」の小さい順に選択するだけでなく、サーバ計算機を所定の規則に基づいて順番に使用するようにしてもよく、また、所定の閾値以下の「負荷」を有するサーバ計算機をランダムに選択するようにしてもよい。

【０２４８】

【発明の効果】以上のように、本発明の分散ファイル管理装置及び分散ファイル管理システムによれば、クライアント計算機からサーバ計算機に対する分散ファイルの作成、参照、または更新の要求に応じて、要求されたサーバ計算機または管理サーバ計算機の各々の管理部で、各サーバ計算機の負荷情報に基づいて部分ファイルの配置を決定するため、特定のサーバ計算機への負荷の集中を回避することができるようになった。

【０２４９】また、他のサーバ計算機へ負荷情報を通知し、また、他のサーバ計算機から通知された外部負荷情報を保持して、他のサーバ計算機の負荷情報に基づいて、部分ファイルの配置を決定するため、特定のサーバ計算機への負荷の集中を回避することができるようになった。

【０２５０】また、部分ファイル毎のアクセス情報、負荷情報及び外部負荷情報に基づいて、移動させる部分ファイルを決定し、他のサーバ計算機へ部分ファイルを移動するため、特定のサーバ計算機の記憶装置への負荷の集中や、各サーバ計算機の記憶装置の容量の不均衡を回避することができるようになった。

【０２５１】また、部分ファイル毎のアクセス状況、負荷情報及び外部負荷情報に基づいて、コピーする部分ファイルを決定し、他のサーバ計算機に部分ファイルをコピーするため、特定のサーバ計算機の記憶装置への負荷の集中を回避することができるようになった。

【０２５２】また、クライアント計算機からの情報や分散ファイルに記録されるデータの種類によって、分散ファイルを分割して作成する部分ファイルのサイズを決定するため、分散ファイルを構成する部分ファイルのサイズを適宜変更することができ、内容的、論理的に関連のあるデータ、例えば、画像１フレーム分のデータなどを複数の部分ファイルに分けて記録することを防止することができるようになった。

【０２５３】また、分散ファイルの部分ファイルを管理する各管理部を１つまたは複数の管理用サーバ計算機に集中させることにより、リソースの重複を最小限に抑えることができるため、コストの増加を抑えることができるようになった。

【図面の簡単な説明】

【図１】本発明の分散ファイル管理システムを示すブロック図である。

【図２】本発明における分散ファイルの構成を示す図である。

【図３】本発明の分散ファイル管理部の分散ファイル作成アルゴリズムを示すフローチャートである。

【図４】本発明における負荷情報テーブルの一例を示す図である。

【図５】本発明の分散ファイル管理テーブルの一例を示す図である。

【図６】本発明の部分ファイル管理テーブルの一例を示す図である。

【図７】本発明の分散ファイル管理部の分散ファイルの参照／更新アルゴリズムを示すフローチャートである。

【図８】本発明の分散ファイル管理システムを示すブロック図である。

【図９】本発明における外部負荷情報テーブルの一例を示す図である。

【図１０】本発明における分散ファイル管理部の分散ファイル作成アルゴリズムを示すフローチャートである。

【図１１】本発明の分散ファイル管理システムを示すブロック図である。

【図１２】本発明における部分ファイルアクセス情報テーブルの一例を示す図である。

【図１３】本発明における部分ファイル管理テーブルの一例を示す図である。

【図１４】本発明における部分ファイル管理テーブルの一例を示す図である。

【図１５】本発明における分散ファイル移動部の動作アルゴリズムを示すフローチャートである。

【図１６】本発明における部分ファイル管理テーブルの一例を示す図である。

【図１７】本発明における部分ファイル管理テーブルの一例を示す図である。

【図１８】本発明における分散ファイル移動部の動作アルゴリズムを示すフローチャートである。

【図１９】本発明における接続情報テーブルの一例を示す図である。

【図２０】本発明の分散ファイル管理システムを示すブロック図である。

【図２１】本発明における分散ファイルコピー部の動作アルゴリズムを示すフローチャートである。

【図２２】本発明における部分ファイル管理テーブルの一例を示す図である。

【図２３】本発明における部分ファイル管理テーブルの一例を示す図である。

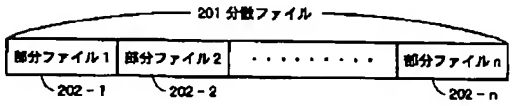
【図２４】本発明における分散ファイルコピー部の動作アルゴリズムを示すフローチャートである。

【図２５】本発明における分散ファイルコピー部の動作アルゴリズムを示すフローチャートである。

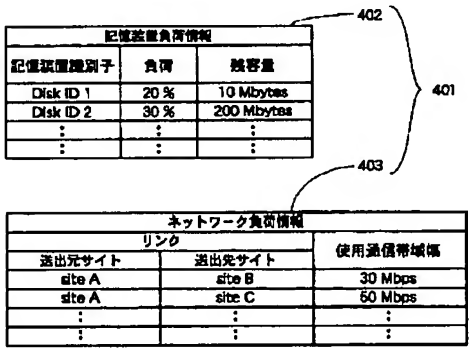
【図２６】本発明の分散ファイル管理システムを示すブ

ロック図である。
【図２７】本発明の分散ファイル管理システムを示すブロック図である。
【図２８】本発明の分散ファイル管理部の分散ファイル作成アルゴリズムを示すフローチャートである。
【図２９】本発明の分散ファイル管理部の分散ファイルの参照／更新アルゴリズムを示すフローチャートである。
【図３０】従来の分散ファイル管理装置を示すブロック図である。
【符号の説明】
１０１　ネットワーク
１０２～１０４、８０２～８０４、１１０２～１１０４、２００２～２００４、２６０２～２６０４、２７０２～２７０４、２７１０　計算機サイト
１０５～１０７、８０５～８０７、１１０５～１１０７、２００５～２００７、２６０５～２６０７、２７０５～２７０７　サーバ計算機
１０８、１０９、１１０　クライアント計算機群
１０８－１～１０８－ｎ　クライアント計算機
１１１、１１６、１２１　部分ファイル管理部
１１２、１１７、１２２、１１１２、１１１７、１１２２、２０１２、２０１７、２０２２、２６１２、２６１７、２６２２、２７１２　分散ファイル管理部
＊１１３、１１８、１２３、２７１３　ネットワークインタフェース
１１４、１１９、１２４、８１４、８１９、８２４、２７１４　状態管理部
１１５、１２０、１２５　記憶装置
１２６ａ～１２６ｆ、２０２－１～２０２－ｎ　部分ファイル
１３１～１３３　内部ネットワーク
４０１　負荷情報テーブル
４０２　記憶装置負荷情報テーブル
４０３　ネットワーク負荷情報テーブル
５０１　分散ファイル管理テーブル
６０１、１３０１、１４０１～１４０３、１６０１～１６０３、１７０１～１７０３、２２０１～２２０３、２３０１～２３０３　部分ファイル管理テーブル
８１１、８１２、８１３　外部状態管理部
９０１　外部負荷情報テーブル
１１３１、１１３２、１１３３　分散ファイル移動部
１２０１　アクセス情報テーブル
１９０１　接続情報テーブル
２０３１、２０３２、２０３３　分散ファイルコピー部
２６３１、２６３２、２６３３　部分ファイルサイズ決定部
＊２７１１　管理サーバ計算機

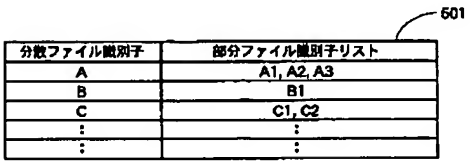
【図２】



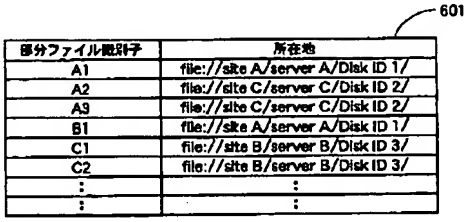
【図４】



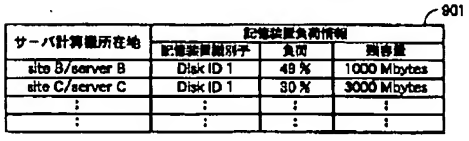
【図５】



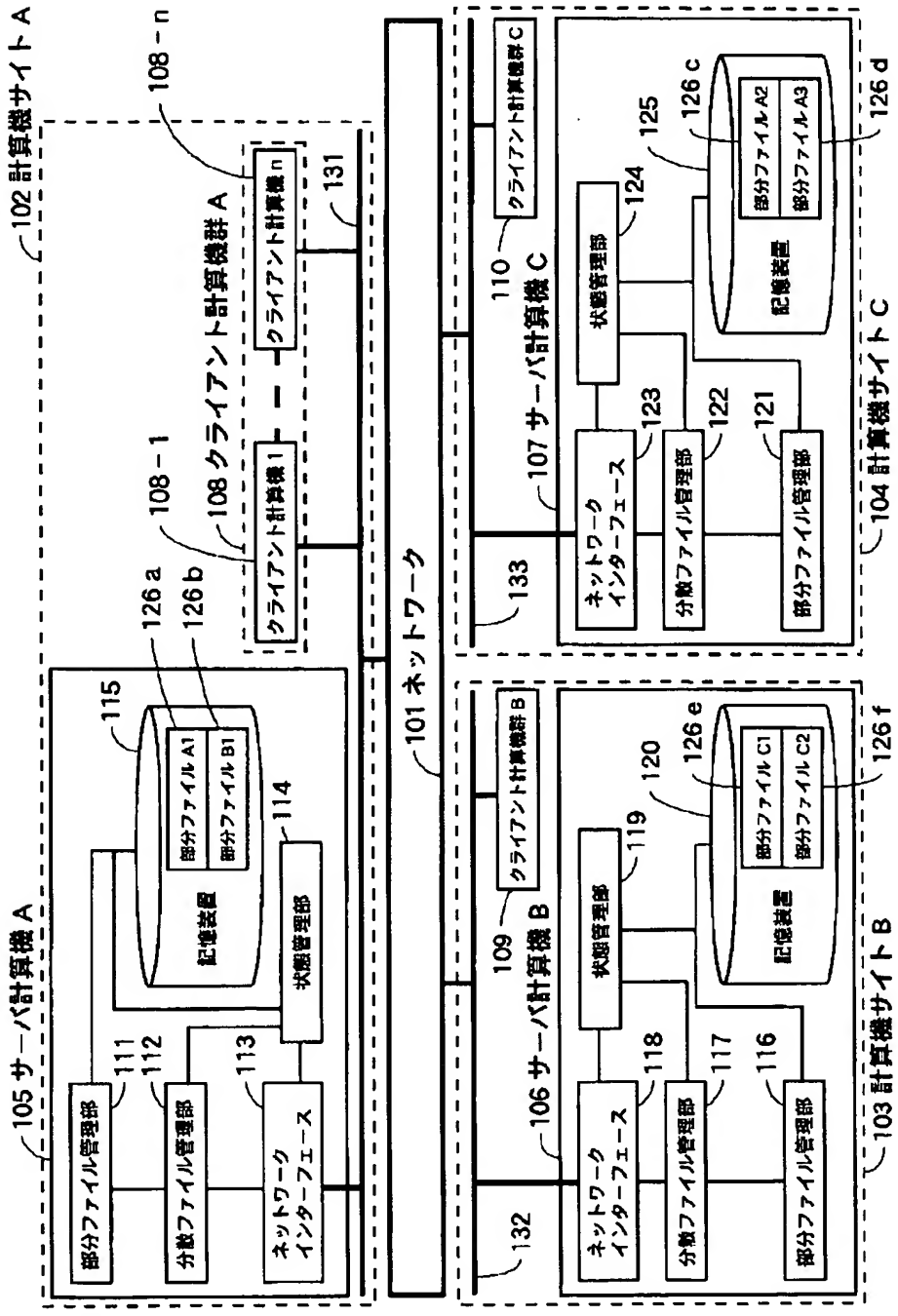
【図６】



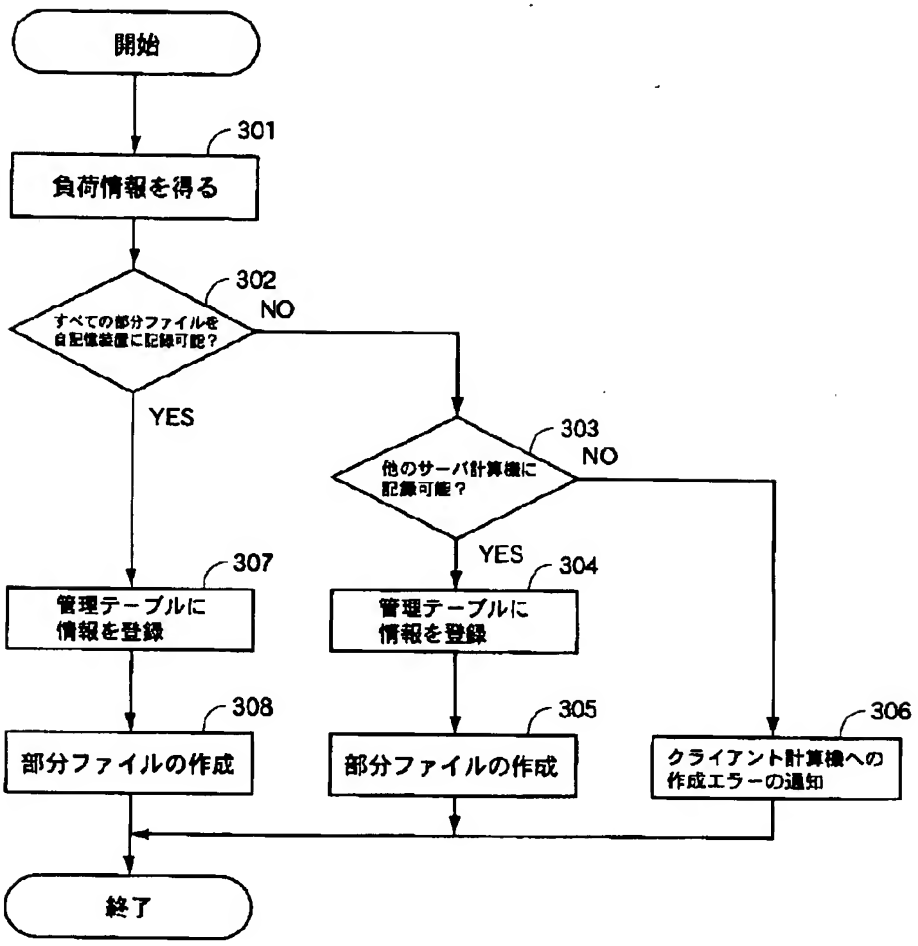
【図９】



【図1】


102 計算機サイトA
105 サーバ計算機A
部分ファイル管理部
111
分散ファイル管理部
112
ネットワークインターフェース
113
状態管理部
114
記憶装置
115
部分ファイルA1
部分ファイルB1
126a
126b
108－1
108－n
108 クライアント計算機群A
クライアント計算機1
クライアント計算機n
131
101 ネットワーク
103 計算機サイトB
106 サーバ計算機B
109
クライアント計算機群B
132
ネットワークインターフェース
118
分散ファイル管理部
117
部分ファイル管理部
116
状態管理部
119
120
記憶装置
126e
126f
部分ファイルC1
部分ファイルC2
104 計算機サイトC
107 サーバ計算機C
110
クライアント計算機群C
133
ネットワークインターフェース
123
分散ファイル管理部
122
部分ファイル管理部
121
状態管理部
124
125
記憶装置
126c
126d
部分ファイルA2
部分ファイルA3

【図３】



【図１２】

1201

部分ファイル識別子	単位時間あたりのアクセス回数	
	アクセス元サイト識別子	アクセス回数
A1	site B	10
A2	site A	2
B1	site C	5
：	：	
：	：	

【図１３】

1301

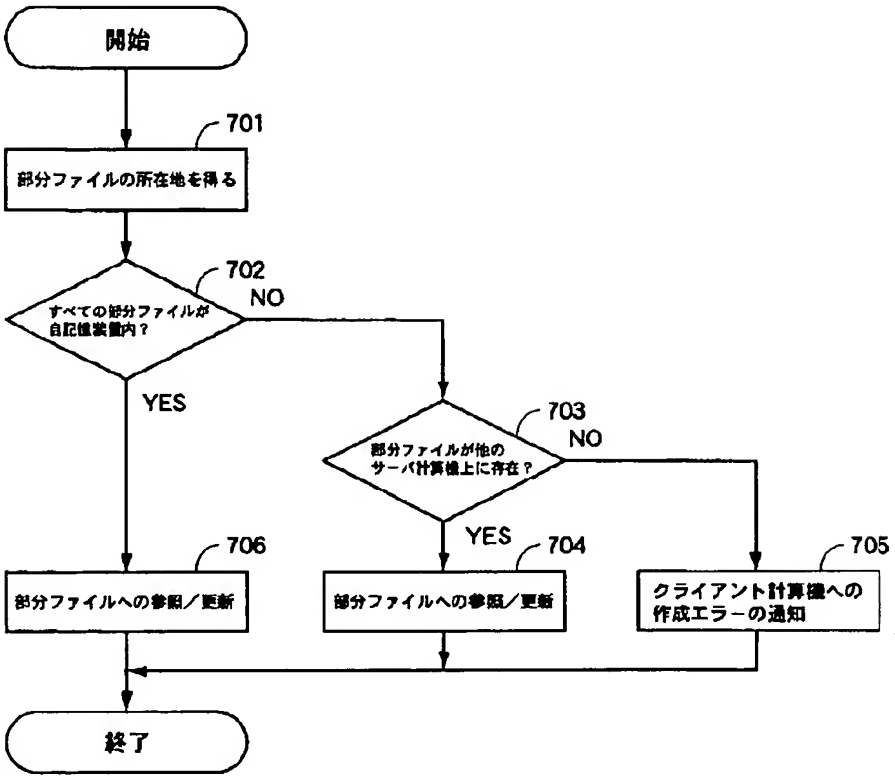
部分ファイル識別子	所在地	オリジナル所在地
A1	file://site A/server A/Disk ID 1/	file://site A/server A/Disk ID 1/
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
B1	file://site A/server A/Disk ID 2/	file://site A/server A/Disk ID 2/
C1	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
：	：	
：	：	

【図１９】

1901

送出元サイト	送出先サイト	通信時間
site A	site B	10 ms
site B	site A	20 ms

【図７】



【図１４】

(A)

1401

部分ファイル識別子	所在地	オリジナル所在地
A1	file://site A/server A/Disk ID 1/	file://site A/server A/Disk ID 1/
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
B1	file://site A/server A/Disk ID 1/	file://site A/server A/Disk ID 1/
C1	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

(B)

1402

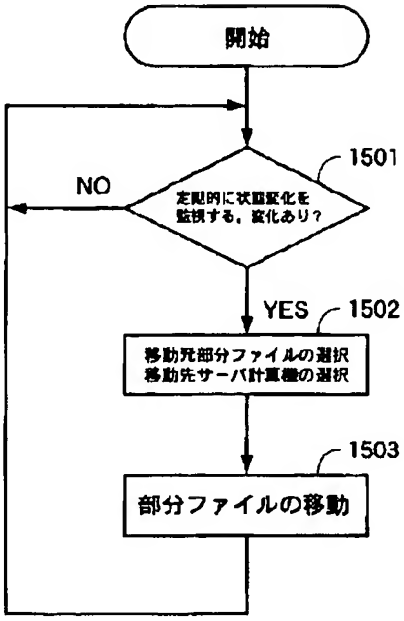
部分ファイル識別子	所在地	オリジナル所在地
C1	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

(C)

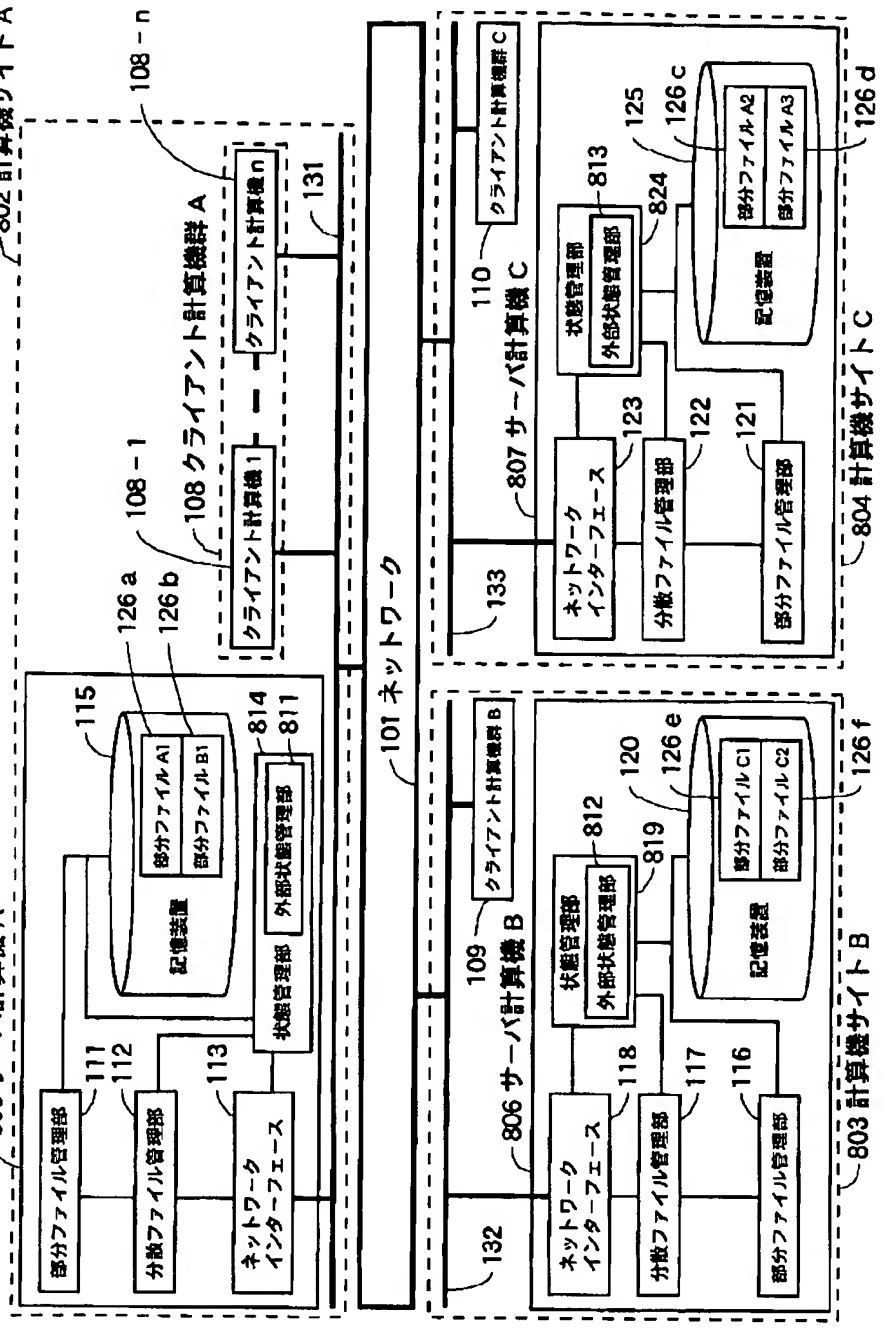
1403

部分ファイル識別子	所在地	オリジナル所在地
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
⋮	⋮	
⋮	⋮	

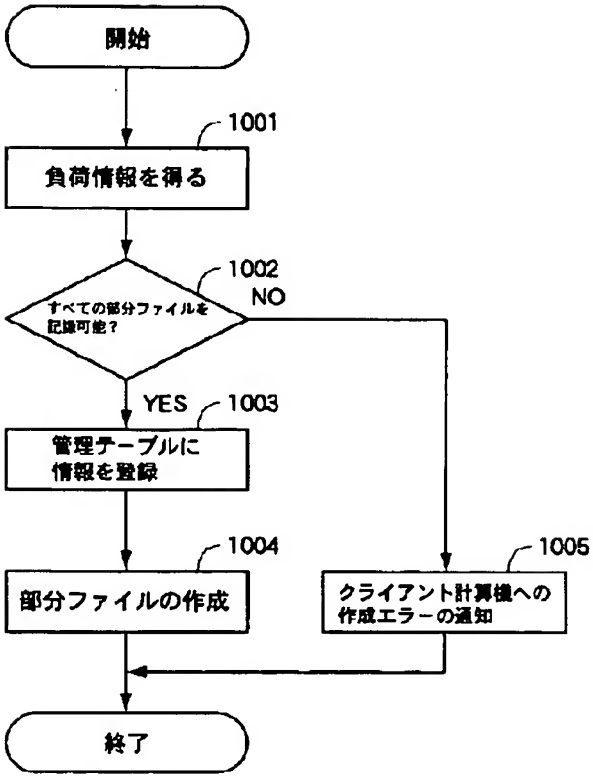
【図１５】



【図8】


802 計算機サイトA
805 サーバ計算機A
806 サーバ計算機B
807 サーバ計算機C
803 計算機サイトB
804 計算機サイトC
101 ネットワーク
108 クライアント計算機群A
108-1
108-n
クライアント計算機1
クライアント計算機n
クライアント計算機群B
クライアント計算機群C
131
132
133
109
110
111
112
113
115
116
117
118
120
121
122
123
125
126a
126b
126c
126d
126e
126f
811
812
813
814
819
824
分散ファイル管理部
部分ファイル管理部
ネットワークインターフェース
記憶装置
部分ファイルA1
部分ファイルB1
部分ファイルA2
部分ファイルA3
部分ファイルC1
部分ファイルC2
状態管理部
外部状態管理部

【図１０】



【図１６】

(A)

1601

部分ファイル識別子	所在地	オリジナル所在地
A1	file://site C/server C/Disk ID 2/	file://site A/server A/Disk ID 1/
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
B1	file://site A/server A/Disk ID 1/	file://site A/server A/Disk ID 1/
C1	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

(B)

1602

部分ファイル識別子	所在地	オリジナル所在地
C1	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

(C)

1603

部分ファイル識別子	所在地	オリジナル所在地
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A1	file://site C/server C/Disk ID 2/	file://site A/server A/Disk ID 1/
⋮	⋮	
⋮	⋮	

【図１７】

(A)

1701

部分ファイル識別子	所在地	オリジナル所在地
A1	file://site B/server B/Disk ID 2/	file://site A/server A/Disk ID 1/
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
B1	file://site A/server A/Disk ID 1/	file://site A/server A/Disk ID 1/
C1	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

(B)

1702

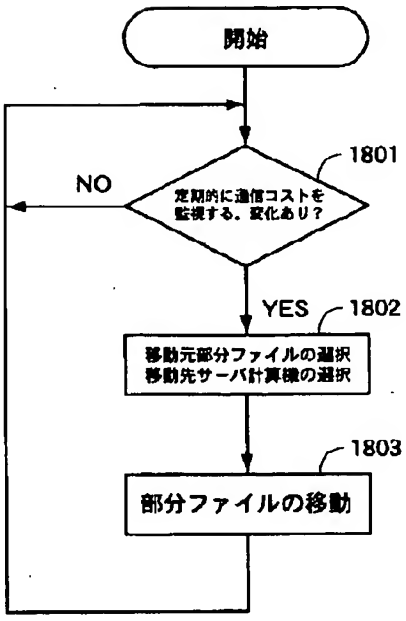
部分ファイル識別子	所在地	オリジナル所在地
C1	file://site C/server C/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
A1	file://site B/server B/Disk ID 2/	file://site A/server A/Disk ID 1/
⋮	⋮	
⋮	⋮	

(C)

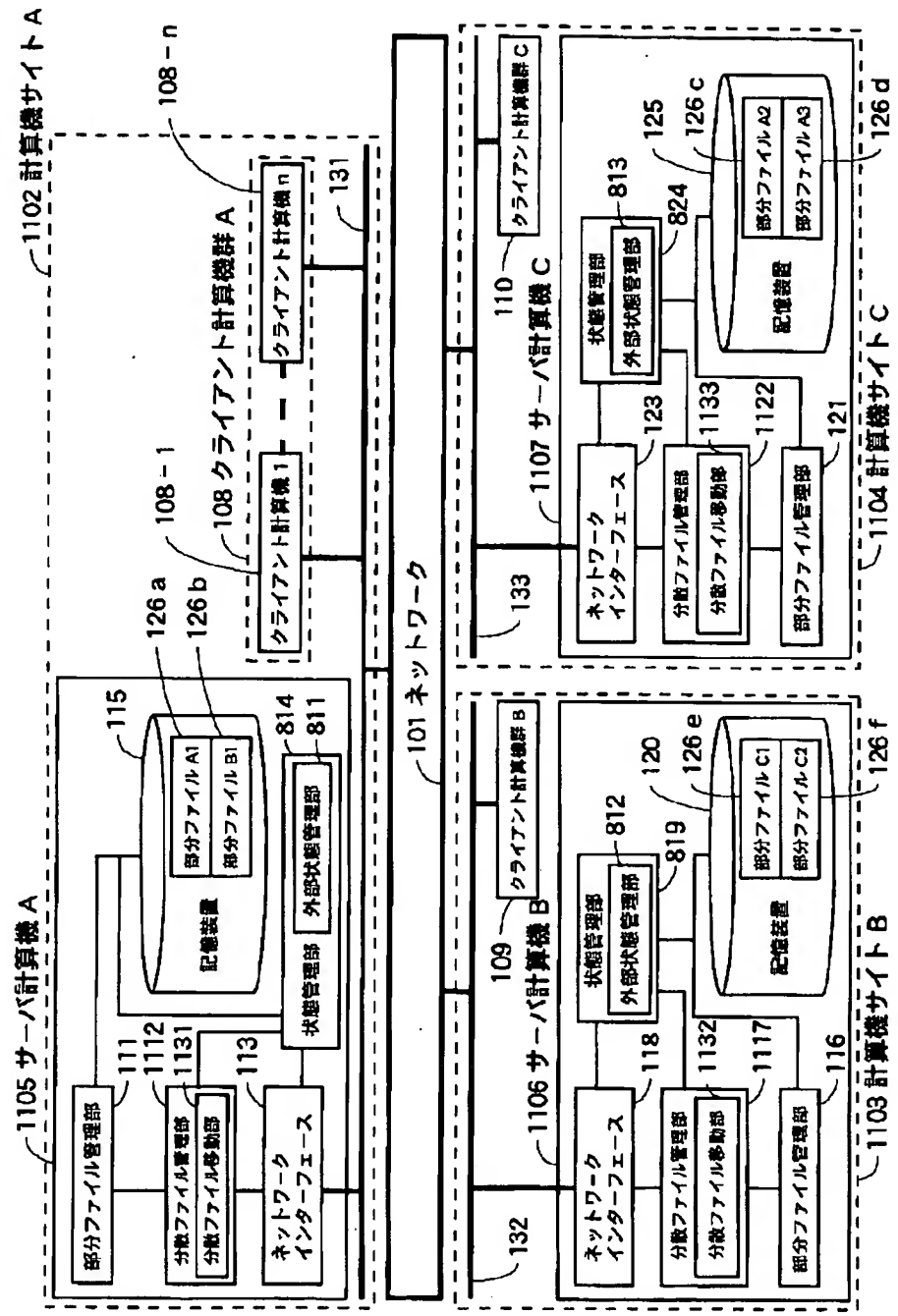
1703

部分ファイル識別子	所在地	オリジナル所在地
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
C1	file://site C/server C/Disk ID 3/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

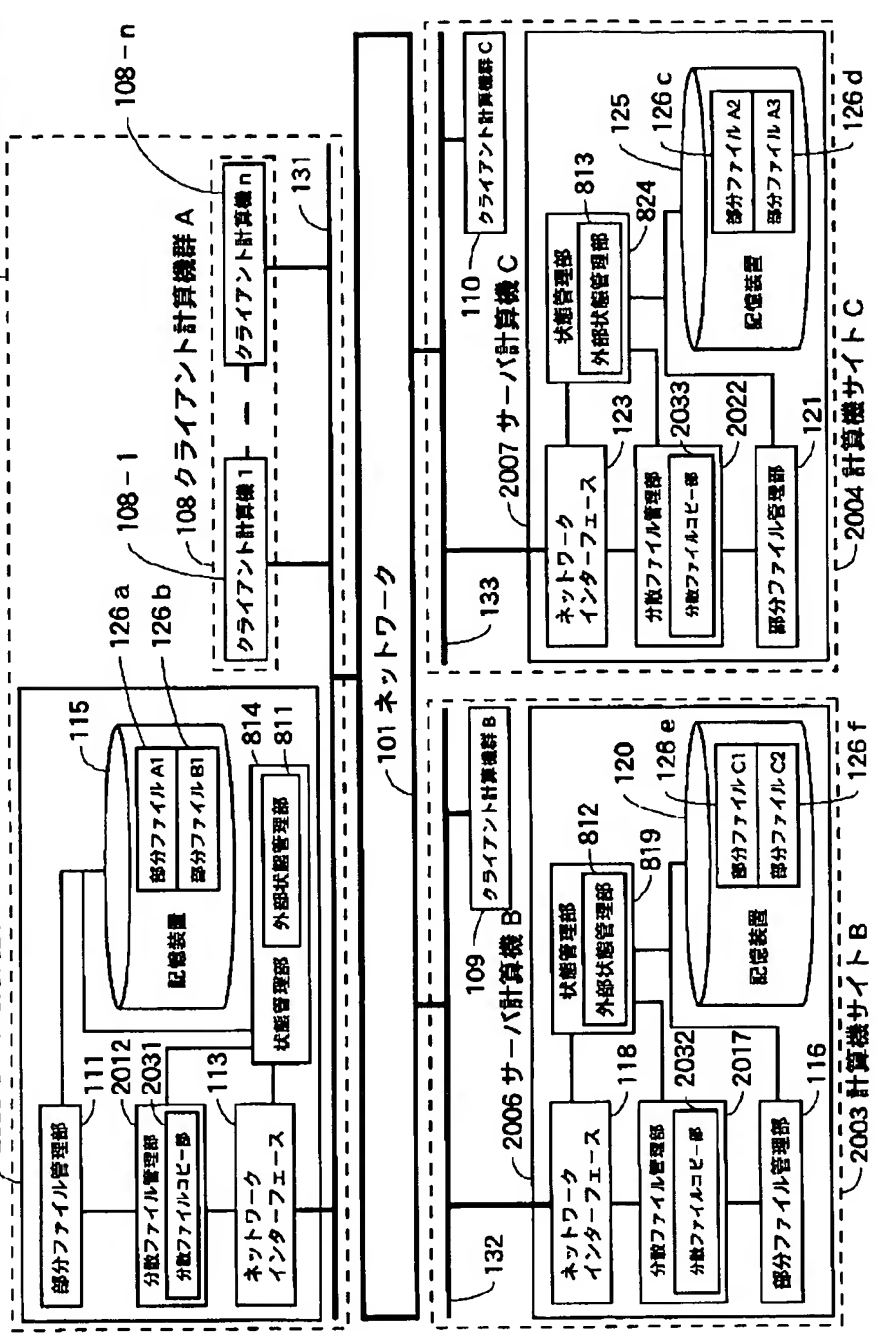
【図１８】



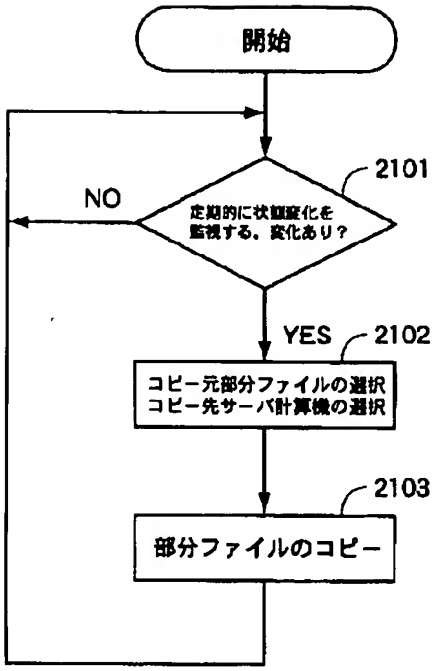
【図１１】


1102 計算機サイトA
1105 サーバ計算機A
部分ファイル管理部
111
分散ファイル管理部
1112
分散ファイル移動部
1131
ネットワークインターフェース
113
115
記憶装置
部分ファイルA1
部分ファイルB1
126a
126b
状態管理部
外部状態管理部
814
811
108-1
108 クライアント計算機群A
クライアント計算機1
108-n
クライアント計算機n
131
101 ネットワーク
1106 サーバ計算機B
109
クライアント計算機群B
132
状態管理部
外部状態管理部
812
819
118
ネットワークインターフェース
分散ファイル管理部
分散ファイル移動部
1132
1117
116
120
126e
126f
記憶装置
部分ファイルC1
部分ファイルC2
部分ファイル管理部
1103 計算機サイトB
1107 サーバ計算機C
110
クライアント計算機群C
133
状態管理部
外部状態管理部
813
824
123
ネットワークインターフェース
分散ファイル管理部
分散ファイル移動部
1133
1122
121
125
126c
126d
記憶装置
部分ファイルA2
部分ファイルA3
部分ファイル管理部
1104 計算機サイトC

【図２０】


2002 計算機サイト A
2005 サーバ計算機 A
2006 サーバ計算機 B
2007 サーバ計算機 C
2003 計算機サイト B
2004 計算機サイト C
101 ネットワーク
108 クライアント計算機群 A
108－1
108－n
クライアント計算機 1
クライアント計算機 n
クライアント計算機群 B
クライアント計算機群 C
部分ファイル管理部
分散ファイル管理部
分散ファイルコピー部
ネットワーク インターフェース
状態管理部
外部状態管理部
記憶装置
部分ファイル A1
部分ファイル B1
部分ファイル C1
部分ファイル C2
部分ファイル A2
部分ファイル A3
111
2012
2031
113
814
811
115
126 a
126 b
131
132
109
812
819
120
126 e
126 f
118
2032
2017
116
110
813
824
125
126 c
126 d
133
123
2033
2022
121

【図２１】



【図２２】

(A)

2201

部分ファイル識別子	所在地	オリジナル所在地
A1	file://site A/server A/Disk ID 1/ file://site C/server C/Disk ID 2/	file://site A/server A/Disk ID 1/
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
B1	file://site A/server A/Disk ID 1/	file://site A/server A/Disk ID 1/
C1	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

(B)

2202

部分ファイル識別子	所在地	オリジナル所在地
C1	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

(C)

2203

部分ファイル識別子	所在地	オリジナル所在地
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A1	file://site C/server C/Disk ID 1/	file://site A/server A/Disk ID 1/
⋮	⋮	
⋮	⋮	

【図２３】

(A)

2301

部分ファイル識別子	所在地	オリジナル所在地
A1	file://site A/server A/Disk ID 1/ file://site C/server C/Disk ID 2/ file://site B/server B/Disk ID 3/	file://site A/server A/Disk ID 1/
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
B1	file://site A/server A/Disk ID 1/	file://site A/server A/Disk ID 1/
C1	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

(B)

2302

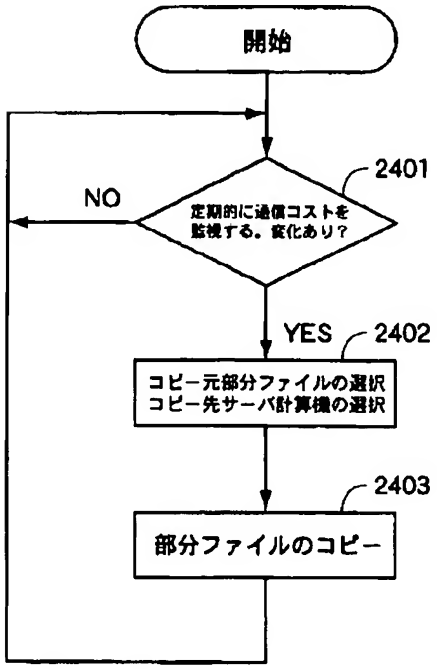
部分ファイル識別子	所在地	オリジナル所在地
C1	file://site B/server B/Disk ID 3/ file://site C/server C/Disk ID 3/	file://site B/server B/Disk ID 3/
C2	file://site B/server B/Disk ID 3/	file://site B/server B/Disk ID 3/
A1	file://site B/server B/Disk ID 2/	file://site A/server A/Disk ID 1/
⋮	⋮	
⋮	⋮	

(C)

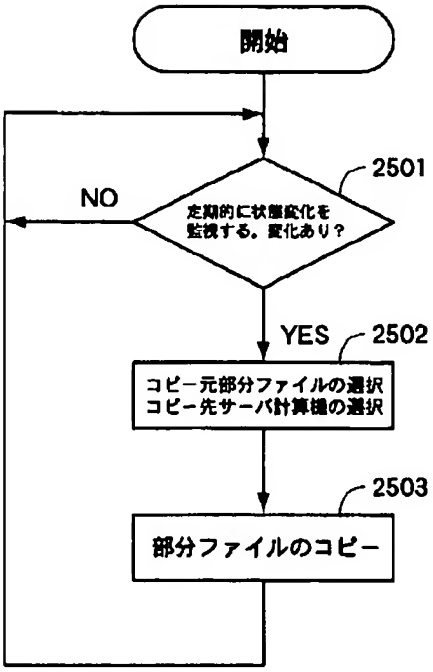
2303

部分ファイル識別子	所在地	オリジナル所在地
A2	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A3	file://site C/server C/Disk ID 2/	file://site C/server C/Disk ID 2/
A1	file://site C/server C/Disk ID 2/	file://site A/server A/Disk ID 1/
C1	file://site C/server C/Disk ID 2/	file://site B/server B/Disk ID 3/
⋮	⋮	
⋮	⋮	

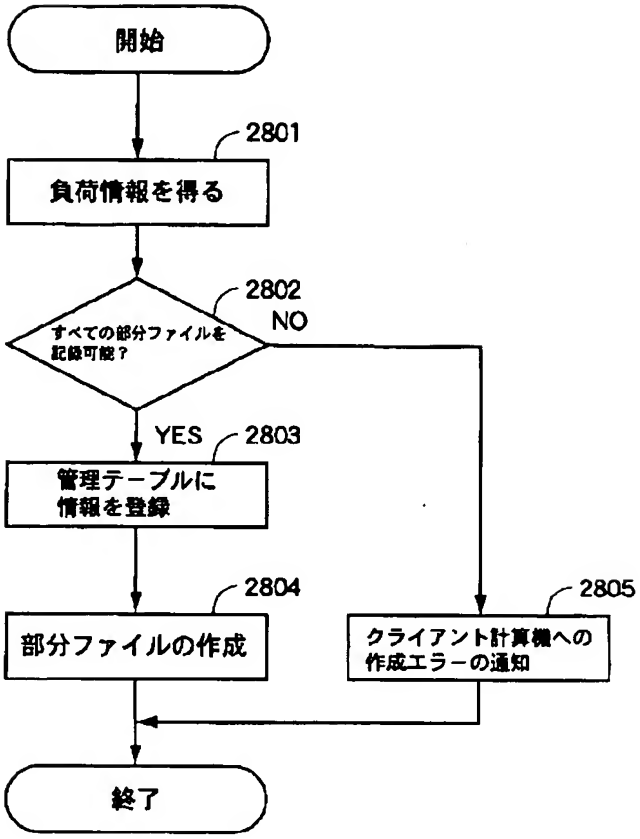
【図２４】



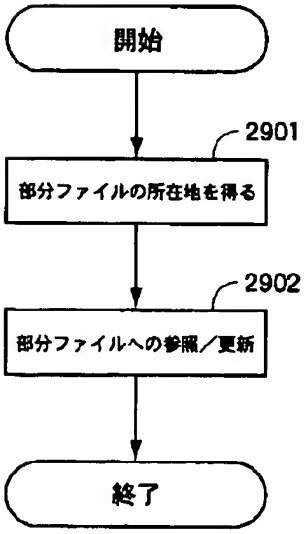
【図２５】


開始
2501
NO
定期的に状態変化を監視する。変化あり？
YES
2502
コピー元部分ファイルの選択
コピー先サーバ計算機の選択
2503
部分ファイルのコピー

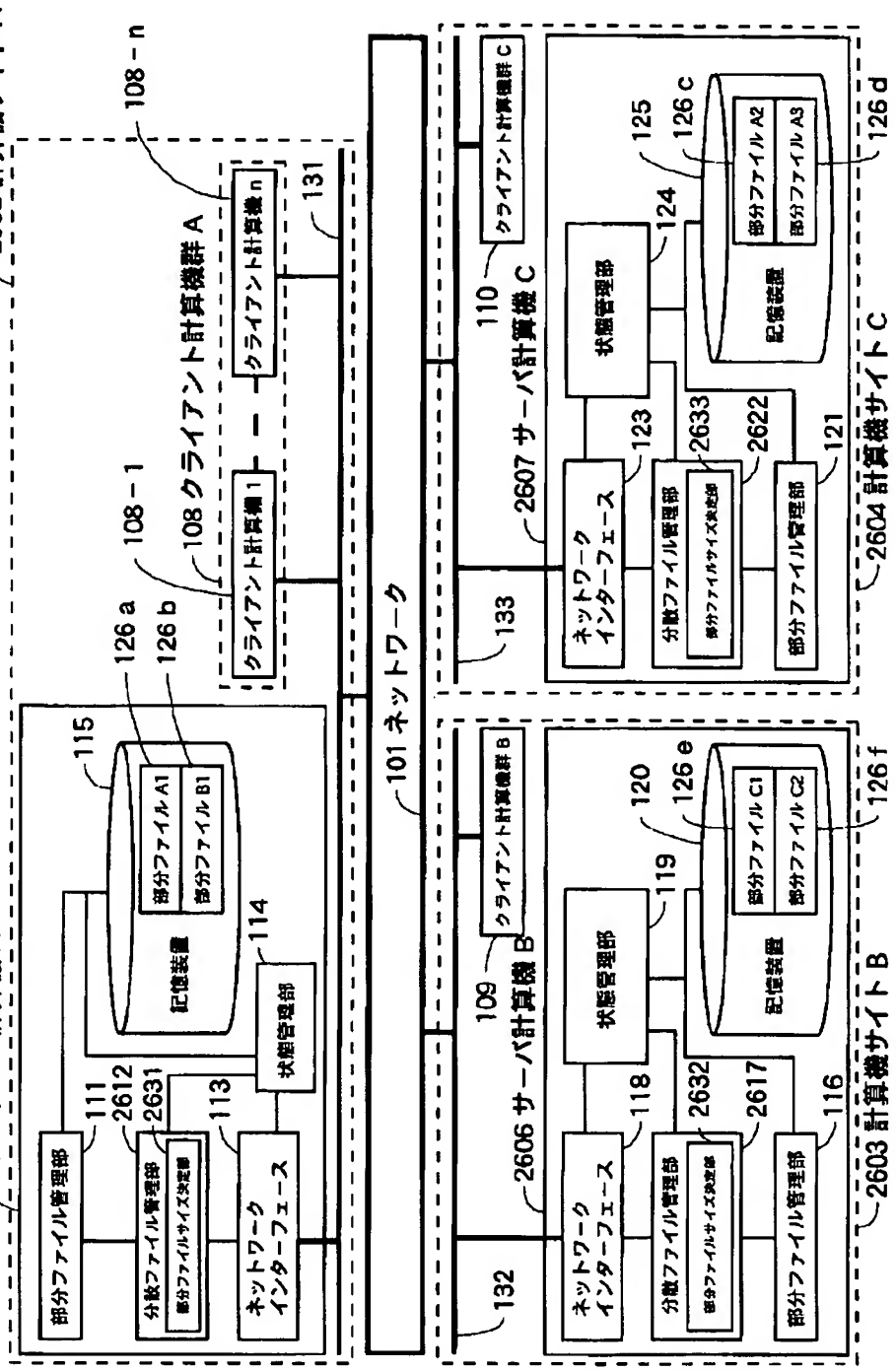
【図２８】


開始
2801
負荷情報を得る
2802
すべての部分ファイルを記録可能？
NO
YES
2803
管理テーブルに情報を登録
2804
部分ファイルの作成
2805
クライアント計算機への作成エラーの通知
終了

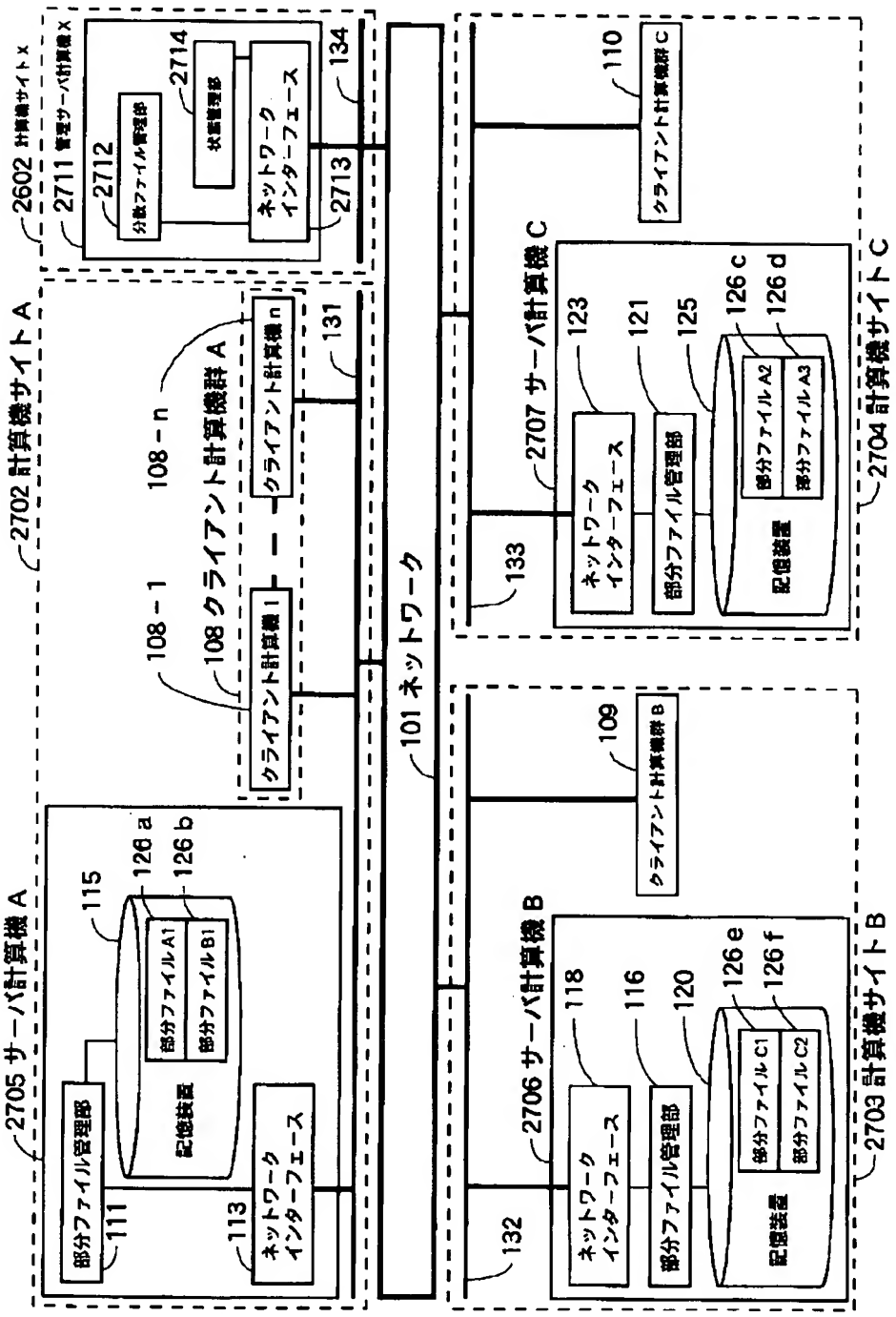
【図２９】


開始
2901
部分ファイルの所在地を得る
2902
部分ファイルへの参照／更新
終了

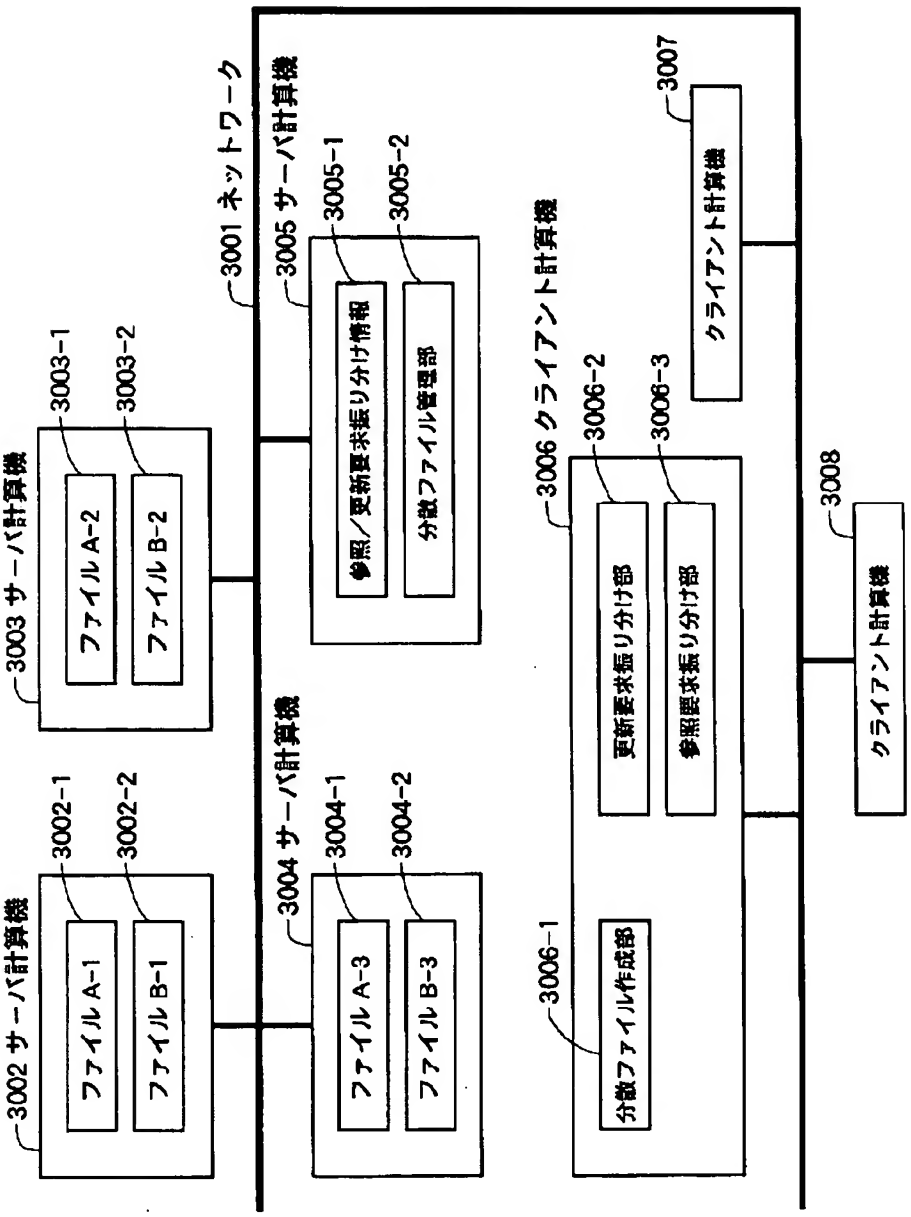
【図26】


2602 計算機サイトA
2605 サーバ計算機A
部分ファイル管理部
111
分散ファイル管理部
2612
部分ファイルサイズ決定部
2631
113
ネットワークインターフェース
状態管理部
114
記憶装置
115
部分ファイルA1
部分ファイルB1
126 a
126 b
108 クライアント計算機群A
108−1
クライアント計算機1
108−n
クライアント計算機n
131
101 ネットワーク
2606 サーバ計算機B
109
クライアント計算機群B
132
ネットワークインターフェース
118
分散ファイル管理部
2632
部分ファイルサイズ決定部
2617
部分ファイル管理部
116
状態管理部
119
120
記憶装置
126 e
部分ファイルC1
部分ファイルC2
126 f
2603 計算機サイトB
2607 サーバ計算機C
110
クライアント計算機群C
133
ネットワークインターフェース
123
分散ファイル管理部
2633
部分ファイルサイズ決定部
2622
部分ファイル管理部
121
状態管理部
124
125
記憶装置
126 c
部分ファイルA2
部分ファイルA3
126 d
2604 計算機サイトC

【図27】


2602 計算機サイトX
2711 管理サーバ計算機X
2712
分散ファイル管理部
2714
状態管理部
ネットワーク
インターフェース
2713
134
2702 計算機サイトA
2705 サーバ計算機A
部分ファイル管理部
111
115
126 a
126 b
部分ファイルA1
部分ファイルB1
記憶装置
113
ネットワーク
インターフェース
108－1
108 クライアント計算機群A
108－n
クライアント計算機1
クライアント計算機n
131
101 ネットワーク
2707 サーバ計算機C
123
121
125
126 c
126 d
部分ファイルA2
部分ファイルA3
記憶装置
ネットワーク
インターフェース
部分ファイル管理部
133
クライアント計算機群C
110
2704 計算機サイトC
2706 サーバ計算機B
118
116
120
126 e
126 f
部分ファイルC1
部分ファイルC2
記憶装置
ネットワーク
インターフェース
部分ファイル管理部
132
クライアント計算機群B
109
2703 計算機サイトB

【図３０】



フロントページの続き

(72)発明者　安河内　龍二
大阪府門真市大字門真1006番地　松下電器産業株式会社内

(72)発明者　田中　則子
大阪府門真市大字門真1006番地　松下電器産業株式会社内

Ｆターム(参考)　5B045 BB49 DD16 GG02 GG09 JJ08
5B082 HA01 HA08